

D. Mitchell

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:43:54 ; Search time 115.93 Seconds
(without alignments)
219.110 Million cell updates/sec

Title: US-09-427-657-2

Perfect score: 2336
Sequence: 1 MHLGFFSVACSLAALLP.....SYSEVRCVPSYWRPQMS 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0601:*

- 1: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SID58/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT:*
- 22: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2336	100.0	419	18	AAW17837 Human foetal liver
2	2336	100.0	419	18	AAW00932 Human Fil4 recepto
3	2336	100.0	419	19	AAW5740 Human vascular end
4	2336	100.0	419	20	AAW6203 Human vascular end
5	2336	100.0	419	21	AAW10648 Human VEGF-C prote
6	2336	100.0	419	21	AAW29048 Human VEGF-C prote
7	2336	100.0	419	21	AAW70749 Human prepro-vascu
8	2336	100.0	419	21	AAW70982 Human vascular end
9	2336	100.0	419	22	AAW57605 Human VEGF-C. Hom
10	2327	99.6	419	18	AAW13833 Human vascular end
11	2326	99.6	419	20	AAW30518 Vascular endothell

12	2326	99.6	419	20	AAW22320 Full length human
13	2326	99.6	419	21	AAW97144 Vascular endothell
14	2326	99.6	419	22	AAW97570 Human VEGF-B prote
15	2325	99.5	419	19	AAW5751 Vascular endothell
16	2323	99.4	419	18	AAW11478 Human vascular end
17	2237	95.8	399	20	AAW66237 Human VEGF-C full
18	2048	87.7	415	18	AAW5742 Mouse vascular end
19	2043	87.5	415	18	AAW09933 Mouse Fil4 recepto
20	1995	85.4	350	20	AAW30519 A truncated vascul
21	1995	85.4	350	20	AAW2321 Truncated human VE
22	1995	85.4	350	21	AAW97145 Truncated vascular
23	1995	85.4	350	22	AAW97577 Human VEGF-2 prote
24	1990	85.2	350	16	AAW82686 Vascular endothell
25	1813	77.6	318	20	AAW08284 Human growth facto
26	1804.5	77.2	418	18	AAW09934 Quail Fil4 recepto
27	1804.5	77.2	418	18	AAW5743 Quail vascular end
28	1761	75.4	307	20	AAW66222 Human VEGF-C trunc
29	1730	74.1	302	20	AAW66223 Human VEGF-C trunc
30	1709	73.2	297	20	AAW66224 Human VEGF-C trunc
31	1676	71.7	292	20	AAW66225 Human VEGF-C trunc
32	704.5	30.2	325	19	AAW53240 Homo sapiens vascu
33	704.5	30.2	325	22	AAW97572 Homo VEGF-D prote
34	704.5	30.2	354	19	AAW49036 Homo zveg12 growt
35	704.5	30.2	354	19	AAW53241 Homo sapiens vascu
36	704.5	30.2	354	19	AAW44293 Homo sapiens vascu
37	704.5	30.2	354	21	AAW10649 Human VEGD protein
38	704.5	30.2	354	21	AAW29049 Human VEGF-D prote
39	704.5	30.2	354	21	AAW70750 Human prepro-vascu
40	704.5	30.2	354	21	AAW70983 Human vascular end
41	704.5	30.2	354	22	AAW97573 Human VEGF-D1 prot
42	704.5	30.2	354	22	AAW37606 Human VEGF-D. Hom
43	703.5	30.1	337	20	AAW08286 Human growth facto
44	703.5	30.1	358	19	AAW53242 Mus musculus vascu
45	703.5	30.1	358	19	AAW44295 Mouse vascular end

ALIGNMENTS

RESULT 1	
AAW17837	standard; Protein: 419 AA.
ID	AAW17837
AC	AAW17837
XX	
DT	13-JAN-1998 (first entry)
XX	
DE	Human foetal liver kinase A binding protein flk-1bp.
XX	
KW	Foetal liver kinase I binding protein; human; flk-1bp; receptor tyrosine kinase; vasculogenesis; angiogenesis;
KW	wound healing; tumour; therapy; antagonist; antibody.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Peptide
FT	1..20 Location/Qualifiers
FT	/label= Sig-peptide
FT	21..419
FT	/label= Mat-protein
FT	/note= "Claim 10"
FT	21..35
FT	/label= "N-terminal
FT	/note= "Claim 9"
XX	
PN	WO9717442-A1.
XX	
PD	15-MAY-1997.
XX	
PF	05-NOV-1996; 96WO-US17564.
XX	
PR	08-NOV-1995; 95US-0554374.
XX	

CC VEGF-C, such as antibodies, can be used to control endothelial
 CC cell proliferation, e.g. lymphangioma or metastatic cancer.
 CC Mouse and quail VEGF-C sequences (see AAM00934-35) have also been
 CC isolated.

XX Sequence 419 AA:

Query Match 100.0%; Score 2336; DB 18; Length 419;
 Best Local Similarity 100.0%; Pred. No. 8.9e-173;
 Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHLGFFSVACSLAAALPGPREAPAAAAAFESGLDSDAEPDAGEATVASKDLEOL 60
 DB 1 mhljffsvacslaaallpgpreapaaaafesglidsdaepdageatvaskdleol 60
 QY 61 RSVSSVDELMVLYPEYKMKYKCOLRKGMQHNREOANLNSTTEETKFAAHYNTLILK 120
 DB 61 rsvssvdelmvllypeykmkykcolrkqgmohnrqoanlnstteetkfaahyntelilk 120
 QY 121 SIDNEMRKTCMPREVCIDVGKEFGVATNTFFKPCVSVYRGGCCNSEGLQCMNTSTSY 180
 DB 121 sidnemrktcmprevcidvgkefgvatntffkpcvsvyrcggcnsqglqcmntstsy 180
 QY 181 LSKTLFETVPLSOGPKPVTTISFANHSCRCMSKLDVYRQVHSIRRSPLATLPCCOAN 240
 DB 181 lsktlfetvplsogpkpvttsfanhscrcmskldvyrvqhsirrsplatlpcqgan 240
 QY 241 KTCPTNYMNNHICRCLAOEDFMSSDAGDSTGDFHICGNKELDETCOCVCRAGLR 300
 DB 241 ktcpnymmnhicrclaqedfmssdagdstgdfhdcgnkeldeetcqvcrraglr 300
 QY 301 PASCGPHKELDRNSOCVCKNKLFPSCGANREPDENTCQCCKRTCPRNQPLNPGKAC 360
 DB 301 pascgphkeldrnsocvccknklfpsgcanreidentcqcckrtcpnqplnpgkac 360
 QY 361 ECTESPQCKLKGKFFHHQTCSCYRRPCTNRQKACEPGFSYSEVCRCPVSYWKRPQMS 419
 DB 361 ectespqckllgkffhnhqtcscyrpctnrqkacepgfsysevcrvpsywkrrpqms 419

RESULT 3

AAW75740 standard; Protein: 419 AA.

AAW75740;

20-NOV-1998 (first entry)

Human vascular endothelial growth factor C protein.

FILE4; vascular endothelial growth factor C; vascular endothelial cell;
 lymphatic endothelial cell; myelopoiesis; angiogenesis; inflammation;
 lymphangiogenesis; oedema; elephantiasis; Milroy's disease.

Homo sapiens.

WO9833917-A1.

06-AUG-1998.

02-FEB-1998; 98WO-US01973.

05-FEB-1997; 97US-0795430.

(LUDWIG) LUDWIG INST CANCER RES.

(UYHE-) UNIV HELSINKI LICENSING LTD.

Alitalo K, Joukov V;

WPI: 1998-437470/37.

N-PSDB; AAV52576.

XX

PT New isolated vascular endothelial growth factor polypeptide(s) -
 PT used to develop products for treating, e.g. cancers, inflammation,
 PT oedema, granulocytopenia or for wound healing or tissue
 PT transplantation

XX Claim 1; Page 112-115; 177pp; English.

The vascular endothelial growth factor C (VEGF-C) polypeptides have
 CC activities affecting growth and migration of vascular endothelial cells,
 CC promoting growth of lymphatic endothelial cells and lymphatic vessels,
 CC increasing vascular permeability, and affecting myelopoiesis. The
 CC products can be used for stimulating angiogenesis, for inhibiting
 CC angiogenesis, for stimulating lymphangiogenesis, treatment or prevention
 CC of inflammation, oedema, elephantiasis, or Milroy's disease. They can
 CC also be used to modulate myelopoiesis, e.g. treating granulocytopenia.
 CC They can also be used for modulating the growth of endothelial cells.
 CC They can also be used to stimulate lymphocyte production and maturation,
 CC and to promote or inhibit trafficking of leucocytes between tissues and
 CC lymphatic vessels or to affect migration in and out of the thymus.

XX Sequence 419 AA:

Query Match 100.0%; Score 2336; DB 19; Length 419;
 Best Local Similarity 100.0%; Pred. No. 8.9e-173;
 Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHLGFFSVACSLAAALPGPREAPAAAAAFESGLDSDAEPDAGEATVASKDLEOL 60
 DB 1 mhljffsvacslaaallpgpreapaaaafesglidsdaepdageatvaskdleol 60
 QY 61 RSVSSVDELMVLYPEYKMKYKCOLRKGMQHNREOANLNSTTEETKFAAHYNTLILK 120
 DB 61 rsvssvdelmvllypeykmkykcolrkqgmohnrqoanlnstteetkfaahyntelilk 120
 QY 121 SIDNEMRKTCMPREVCIDVGKEFGVATNTFFKPCVSVYRGGCCNSEGLQCMNTSTSY 180
 DB 121 sidnemrktcmprevcidvgkefgvatntffkpcvsvyrcggcnsqglqcmntstsy 180
 QY 181 LSKTLFETVPLSOGPKPVTTISFANHSCRCMSKLDVYRQVHSIRRSPLATLPCCOAN 240
 DB 181 lsktlfetvplsogpkpvttsfanhscrcmskldvyrvqhsirrsplatlpcqgan 240
 QY 241 KTCPTNYMNNHICRCLAOEDFMSSDAGDSTGDFHICGNKELDETCOCVCRAGLR 300
 DB 241 ktcpnymmnhicrclaqedfmssdagdstgdfhdcgnkeldeetcqvcrraglr 300
 QY 301 PASCGPHKELDRNSOCVCKNKLFPSCGANREPDENTCQCCKRTCPRNQPLNPGKAC 360
 DB 301 pascgphkeldrnsocvccknklfpsgcanreidentcqcckrtcpnqplnpgkac 360
 QY 361 ECTESPQCKLKGKFFHHQTCSCYRRPCTNRQKACEPGFSYSEVCRCPVSYWKRPQMS 419
 DB 361 ectespqckllgkffhnhqtcscyrpctnrqkacepgfsysevcrvpsywkrrpqms 419

RESULT 4

AAW86203 standard; Protein: 419 AA.

AAW86203;

16-FEB-1999 (first entry)

Human vascular endothelial growth factor (VEGF)-C sequence.

VEGF; VRF; vascular endothelial growth factor; VEGF-related protein;
 KW recombinant; truncated; gene therapy; angiogenesis; cardiac ischaemia;
 KW coronary; collateral vessel development; cell growth; migration; heart;
 KW lower limb ischaemia; stroke; peripheral vascular disease; intestine;
 KW wound healing; skin; vascular permeability.

Homo sapiens.

XX MO9849300-A2.
 XX 05-NOV-1998.
 XX 20-APR-1998; 98WO-US07801.
 XX 25-APR-1997; 97US-0842984.
 XX (COLL-) COLLATERAL THERAPEUTICS.
 XX Bohlen P;
 XX WPI: 1999-009426/01.
 XX
 PT New truncated vascular endothelial growth factor-related protein
 PT subunits - lack part of the N-terminal sequence, used to stimulate
 PT angiogenesis, e.g. for treating heart disease and ischaemia
 XX
 PS Disclosure; Fig 1; 113pp; English.
 XX
 CC This represents the amino acid sequence of human vascular endothelial
 CC growth factor (VEGF)-C protein. The invention provides truncated VRF
 CC (VEGF-related protein) subunits that have at least one amino acid
 CC N-terminal to the first Cys of the core sequence deleted. Host cells
 CC transformed or transfected with expression vectors containing nucleic
 CC acids encoding the truncated VRF subunits are used to produce the
 CC truncated proteins recombinantly. The truncated VRF subunits, optionally
 CC expressed from gene therapy vectors, have in vivo and in vitro angiogenic
 CC activity and are used to stimulate angiogenesis, particularly coronary
 CC collateral vessel development in cases of cardiac ischaemia; to stimulate
 CC endothelial cell growth and migration in vitro; to treat heart disease;
 CC to treat ischaemia (e.g. cardiac, chronic coronary or chronic lower limb
 CC ischaemia; stroke and peripheral vascular disease); to promote healing of
 CC wounds (of skin or intestines), and to increase vascular permeability.
 XX
 SQ Sequence 419 AA;

Query Match 100.0%; Score 2336; DB 20; Length 419;
 Best Local Similarity 100.0%; Pred. No. 8.9e-173;
 Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHLGFFSVACSLAALPGPREAPAAAAAFESGLDSDAEPDAGEATAYASKDLEOL 60
 DB 1 mhlglffsvacslaaallpgpreapaaaaafesgldsdapeadageatayaskdleeql 60
 QY 61 RSVSSVDELMTVLYPEYKMYKCOLRKGMQHNRRQANLNSTREETIKFAAHYNTIELIK 120
 DB 61 rsvssvdelmtvlypeykmkycqlrkqgwqhnrqanlnstteetlkfaaahyntelilk 120
 QY 121 SIDNEWRKTCQCPREKCIQVNGKEFGVATNMFPCVSVYRCGGCCNSGLCOMNTSTSY 180
 DB 121 sidnewrktcqcmprckciqvgkefyvatntffkppcvsvyrcggccnsqldqmntstsy 180
 QY 181 LSKTLEFETVPLSQGKRPVTISFANHTSCRCMSKLDVYRQVSHIIRSLPATLPQQAAN 240
 DB 181 lsktlefeti vplsqgk r pvtisfanhtscrcmskldvyrvqshilrrslpatlpqqaan 240
 QY 241 KTCPTNYMNNHTICRCLAQEDPFESSDAGDDSTDFHDIICGPNKLEDETCCQCVCRAGIR 300
 DB 241 ktcptnymmnhitcrlaqedpfessdagddstdfhdiicgpnkledetccqcvcraglr 300
 QY 301 PASCGPHKELDNSQOCVCKNLFPSSOGANRPEDNTCCQCVCKRCPNPNOLNPGKAC 360
 DB 301 pascgphkel d nscqvcvckn l f pssogane r pedntccqvcvckrcpnngl npgkac 360
 QY 361 ECTESPOKCLGKRFHHTOTCSYRPPCTNRQACPEGFSEYEEVCRVPSYWKRPQMS 419
 DB 361 ectespqkcllgkrfhhtotcsyrppctnrqacepgfsyseevcrvpsywkrlpqms 419

RESULT 5

AAB10648
 ID AAB10648 standard; Protein; 419 AA.
 XX
 AC AAB10648;
 XX
 DT 19-JAN-2001 (first entry)
 XX
 DE Human VEGC protein.
 XX
 KW VEGF-X; vascular endothelial growth factor; human; vulnery; cytosatic;
 KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
 KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
 KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
 KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
 KW venous sore; diabetic ulcer; burns; skin graft growth; VEGC.
 XX
 OS Homo sapiens.
 XX
 PN WO200037641-A2.
 XX
 PD 29-JUN-2000.
 XX
 PF 21-DEC-1999; 99WO-US30503.
 XX
 PR 22-DEC-1998; 98GB-0028377.
 PR 18-MAR-1999; 99US-0124967.
 PR 08-NOV-1999; 99US-0164131.
 XX
 FA (JANNC) JANSSEN PHARM NV.
 XX
 PI Gordon RD, Sprengel JJ, Yon JR, Dijkmans JH, Goslowska A;
 PI Dhanaraj SN, Xu J;
 DR WPI: 2000-442669/38.
 XX
 PT New vascular endothelial growth factor protein, useful for treating or
 PT preventing diseases associated with inappropriate angiogenesis activity
 PT such as cancer, rheumatoid arthritis, psoriasis and wounds
 XX
 PS Disclosure; Fig 11; 127pp; English.

CC This invention describes a novel vascular endothelial growth factor-X
 CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
 CC vulnery, cytosatic, antirheumatic, antiarthritic, antipsoriatic and
 CC antidiabetic activity and acts as an angiogenesis and vascularization
 CC regulator. An antisense molecule of the invention is useful for treating
 CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
 CC retinopathy by inhibiting angiogenic activity or inappropriate
 CC vascularization including formation and proliferation of new blood
 CC vessels, growth and development of tissues, tissue regeneration and organ
 CC and tissue repair in a subject. The products of the invention are useful
 CC for preparing medicaments for treating wounds such as dermal ulcers,
 CC pressure sores, venous sores, diabetic ulcers and burns and to promote
 CC skin graft growth, tissue repair, proliferation of new blood vessels,
 CC tissue regeneration and organ repair by promoting angiogenic activity or
 CC vascularization. This sequence represents the human VEGC protein used
 CC to illustrate the method of the invention.

SQ Sequence 419 AA;
 Query Match 100.0%; Score 2336; DB 21; Length 419;
 Best Local Similarity 100.0%; Pred. No. 8.9e-173;
 Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHLGFFSVACSLAALPGPREAPAAAAAFESGLDSDAEPDAGEATAYASKDLEOL 60
 DB 1 mhlglffsvacslaaallpgpreapaaaaafesgldsdapeadageatayaskdleeql 60
 QY 61 RSVSSVDELMTVLYPEYKMYKCOLRKGMQHNRRQANLNSTREETIKFAAHYNTIELIK 120
 DB 61 rsvssvdelmtvlypeykmkycqlrkqgwqhnrqanlnstteetlkfaaahyntelilk 120

Oy 121 SIDNEMRKTCMPREVCIDVGKEFGVATNTPPKPCVSVYRGGCCNSEGLQCMNTSTSY 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 sidnemrktcmprvcidvgkefgvatntfkkppcvsvyrcgscnsegjqcmntstsy 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Oy 181 LSKTLFEITVPLSGPKPVTTISFANHTSCRCMSKLDVYROYHSIIRSLPATLPQCOQAN 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 181 lsktlfeitvplsgpkpvtisfanhtscrcmskldvyroyhsilirslpatlppcqgan 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Oy 241 KTCPTNYMNMNHICRCLAOEDFMFSSDAGDDSTGDFHDCGPNKELDETCQVCYRAGLR 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 241 ktcpnyymnmnhicrclagdfmfssdagddstdgfhdcgpnkeldeetcqvcrraglr 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Oy 301 PASCGRPHKELDRNSGCQCYCKNKLFPSCGAGNREFEENTCQCYCKRTCPRNOPINPGKAC 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 301 pascgrphkeldrnscgcvccknklfpscganrefeentcgcckrtcpnrpnpgkac 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Oy 361 ECTESPOKCLKGKKFHHQTCSCYRRPCTNRQKACEPGEFSYSEEVCRCPVSYWKRPQMS 419
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 361 ectespokcllkgkffhqtscyrpctnrqkacepgfsyseevcrvcpsywkrrpms 419
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 6

AAB29048
 ID AAB29048 standard; Protein: 419 AA.
 AC AAB29048;

XX 31-JAN-2001 (first entry)

DT Human VEGF-C protein sequence.

DE Human VEGF-C protein sequence.

XX Human; Flt4: fms-like tyrosine kinase 4; Lymphoedema;

KW vascular endothelial growth factor receptor 3; VEGFR-3;

KM Milroy-Nonne syndrome; Lymphoedema praecox; VEGF-C;

XX Homo sapiens.

OS WO200058511-A1.

XX 05-OCT-2000.

PD 26-MAR-1999; 99WO-US06133.

PF 26-MAR-1999; 99WO-US06133.

PR 26-MAR-1999; 99WO-US06133.

XX (LUDM-) LUDMIG INST CANCER RES.

PA (UYHE-) UNIV HELSINKI LICENSING LTD OY.

KM (UYPI-) UNIV PITTSBURGH.

XX Ferrell RE, Altalo K, Fliegold DN, Karkkainen M;

PI WPI: 2000-679298/66.

XX N-PSDB: AAC62406.

DR Screening a human subject for increased risk of developing a lymphatic

XX disorder, comprises assaying a nucleic acid to determine a mutation

PT altering the sequence of a vascular endothelial growth factor

PT receptor-3 -

XX Disclosure: Page 60-61; 76pp; English.

XX The present sequence is the protein sequence for the human vascular

CC endothelial growth factor C (VEGF-C). It was used to demonstrate the

CC methods of the invention, which involve the screening of individuals to

CC determine which vascular endothelial growth factor receptor 3 (VEGFR-3,

CC also known as Flt4 or fms-like tyrosine kinase 4) alleles they possess

CC and thus their likelihood of developing hereditary lymphoedema.

CC Conditions associated with lymphoedema include Milroy-Nonne syndrome,

CC which is early onset lymphoedema and lymphoedema praecox, which is late

CC onset.

XX Sequence 419 AA;

Query Match 100.0%; Score 2336; DB 21; Length 419;
 Best Local Similarity 100.0%; Pred. No. 8,9e-173;
 Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MHLLGEFVACSLAAALPGPREAPAAAAAFESGLDSDAPDAGEAVAYASKLEOL 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 mhllgfvacsllaaalpgpreapaaaafesjldsdapdagatayaskleol 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Oy 61 RSVSVDELMTVLYPEYKMYKQCLRKGMOWHNRQANLNSTEETIKFAAAHYMTETLK 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 rsvsvdelmtvlypeykmykqclrkgmowhnrqanlnsteetikfaaahytmteilk 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Oy 121 SIDNEMRKTCMPREVCIDVGKEFGVATNTPPKPCVSVYRGGCCNSEGLQCMNTSTSY 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 sidnemrktcmprvcidvgkefgvatntfkkppcvsvyrcgscnsegjqcmntstsy 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Oy 181 LSKTLFEITVPLSGPKPVTTISFANHTSCRCMSKLDVYROYHSIIRSLPATLPQCOQAN 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 181 lsktlfeitvplsgpkpvtisfanhtscrcmskldvyroyhsilirslpatlppcqgan 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Oy 241 KTCPTNYMNMNHICRCLAOEDFMFSSDAGDDSTGDFHDCGPNKELDETCQVCYRAGLR 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 241 ktcpnyymnmnhicrclagdfmfssdagddstdgfhdcgpnkeldeetcqvcrraglr 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Oy 301 PASCGRPHKELDRNSGCQCYCKNKLFPSCGAGNREFEENTCQCYCKRTCPRNOPINPGKAC 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 301 pascgrphkeldrnscgcvccknklfpscganrefeentcgcckrtcpnrpnpgkac 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Oy 361 ECTESPOKCLKGKKFHHQTCSCYRRPCTNRQKACEPGEFSYSEEVCRCPVSYWKRPQMS 419
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 361 ectespokcllkgkffhqtscyrpctnrqkacepgfsyseevcrvcpsywkrrpms 419
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 7

AAY70749
 ID AAY70749 standard; Protein: 419 AA.

AC AAY70749;

XX 17-AUG-2000 (first entry)

DT Human prepro-vascular endothelial growth factor C.

DE Human; receptor tyrosine kinase; RTK; Flt4: fms-like tyrosine kinase 4;

KW VEGFR-3; Vascular endothelial growth factor receptor-3; Chromosome 5q35;

KM cytoskeletal; tumour imaging; anti-tumour therapy; treatment; diagnosis;

KM neoplastic disease; lymphoma; carcinoma; breast; squamous cell; melanoma;

KM sarcoma; malignancy; VEGF-C; Vascular endothelial growth factor C.

XX Homo sapiens.

OS Location/Qualifiers

FH 1..31

FT Peptide

FT /label= Signal_peptide

FT 32..103

FT /label= N-terminal_peptide

FT /note= "cleavage of this peptide from partially processed

FT VEGF-C produces a fully processed mature form of VEGF-C

FT of 21-23 kD which has high affinity to VEGFR-2"

FT 104..227

FT Protein

FT /label= Mature_VEGF-C

FT 228..419

FT /label= C-terminal_peptide

FT /note= "Has a pattern of spaced cysteine residues

FT reminiscent of a Balbiani ring 3 protein (BR3P) sequence;

FT cleavage of signal peptide and the C-terminal

FT peptide produces a partially processed form of VEGF-C of

FT about 29 kD which has high affinity to Flt4 (VEGFR-3)"

FT 113..213

FT Binding-site

FT /note= "binds and stimulates VEGF-C receptors; Cys

FT at position 156 is essential for VEGFR-2 binding and at

FT	Region	165 is essential for VEGFR-2 and VEGFR-3 binding"
FT		131..211
FT		/note="Important for VEGF-C activity"
PN	WO200021560-A1.	
XX	20-Apr-2000.	
PD	08-OCT-1999;	99WO-US23525.
XX	08-OCT-1999;	99MO-US23525.
PR	09-OCT-1998;	98US-0169079.
XX	(LUDW-) LUDWIG INST CANCER RES.	
PA	(UYHE-) UNIV HELSINKI LICENSING LTD OY.	
XX		
PI	Alltalo K, Kaipainen A, Vaittola R, Jussila L;	
XX	WPI: 2000-317850/27.	
DR		
XX	Treating neoplastic diseases such as lymphoma, carcinomas, melanomas	
PT	and sarcomas, involves administering a compound capable of inhibiting	
XX	binding of ligand proteins to fms-like tyrosine kinase-4 receptor	
PT		
XX	Example 15-17, Page 140-142; 148pp: English.	
PS		
XX	The patent discloses a method to treat neoplastic disease characterised	
XX	by expression of fms-like tyrosine kinase 4 (Flt4) receptor (also	
CC	referred as vascular endothelial growth factor receptor-3, VEGFR-3) in	
CC	endothelial cells of blood vessels adjacent to malignant neoplasm. The	
CC	method involves administering a compound that inhibits binding of a	
CC	ligand to Flt4 thereby inhibiting Flt4 mediated proliferation of vascular	
CC	endothelial cells. The compound is useful for treating neoplastic disease	
CC	such as breast carcinomas, squamous cell carcinomas, lymphomas, melanomas	
CC	and sarcomas. Flt4 receptor tyrosine kinase binding compounds can be used	
CC	for manufacturing medicament useful for diagnostic screening, imaging and	
CC	treatment of malignancies characterised by Flt4-expressing blood cells.	
CC	The Flt4 gene maps to chromosomal region 5q35 and is expressed as 5.8 kb	
CC	and 4.5 kb mRNAs which differ in their 3' sequences and are	
CC	differentially expressed in HEL and DAMI cell lines. Flt4	
CC	belongs to a subfamily of class III receptor tyrosine kinases (RTKs).	
CC	It is used as a target for tumour imaging and anti-tumour therapy.	
CC	The present sequence is a human prepro-vascular endothelial growth	
CC	factor C (VEGF-C), a specific example of Flt4 binding compound.	
XX		
CC		
XX	Sequence 419 AA:	
SO		
Query Match	100.0%: Score 2336; DB 21; Length 419;	
Best Local Similarity	100.0%: Pred. No. 8,9e-173;	
Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY 1	MHLGFEFVSACSLAALPLGPREPAPAAAFAESGLDLSDAEPDAGEATVAVSKDLEOL	60
Db 1	mhllgffvsvacsllaaalhprrpaaapaaafesglldlsdaepdageatayaskdlleql	60
QY 61	RSVSSVDELMTAVLYPEYWKMYKCOLRKSGOMNRRDOANINSTEETIKRAAAHYNETLK	120
Db 61	rsvssvdelmtcvlypeywkmykcqlrksgwqhmrqeanlnsteetikaahaynetllk	120
QY 121	SIDNEMRKTQCMPREVCIDVGKEFGVATNTEFKPCCSVYRKGAGCCNSEGLDCMNTSTSY	180
Db 121	sidnemrktqcmprcvcidvgkefvgatnltfkpccsvyrcgcgcnseglqcmntstay	180
QY 181	LSKTLFEETVPLPSOGPKVNTSFAMHTSGRCMSKLDVYRQVMSITRSLPALPLPQCAAN	240
Db 181	lsktlfeetvplpsgpkpvtlslfahntscrcmskldvyrvqvsilrtslpalplpqqaan	240
QY 241	KTCPTNVMNMNHICQLAQEDFMFSSDAGDDSTDFHHCIGNKLEDETECCQCVRAGR	300
Db 241	ktcptnvmnmnhicrcclaqedfmfssdagddstdfhndicgnkelledetccgcvtcragr	300
QY 301	PASGCPHKELDNRSQCVCYCKNKLFPSSCGANREDENTCCQVCKRTCPRNQPLNPGKAC	360
Db 301	pasgcpfhkeldnrsqcvcycknklfpsscganreudentccqvcvckrtcprnqplnpgkac	360

Accession	Protein Name	Location/Qualifiers
D00001	301 pascgpbkeldrnscgvcknklfpgcsganrefdentcgvckrtcprnglndpnykcaoc	360
D00002	361 ECTSPQKCLLKGGKFFHHQTCSCTRPCTNRKQACGPGSYSEVNCRCVPSYWKRPQMS	419
D00003	361 ectespgkcllkqgkffhhqtcscyrpctnrqkacepgfsyseevrcvpsywxkripqms	419
RESULT 8		
AA070982	AA070982 standard: Protein; 419 AA.	
AA070982:		
09-AUG-2000	(first entry)	
Human	vascular endothelial growth factor (VEGF)-C protein.	
Vascular	endothelial growth factor-C; VEGF; human; re-endothelialisation;	
vascular	endothelial growth factor receptor; VEGFR; vascular trauma;	
blood	vessel; cardiovascular surgery; anti-restenosis agent; prevention;	
restenosis;	stenosis; percutaneous transluminal coronary angioplasty.	
Homo sapiens.		
Key	Location/Qualifiers	
Peptide	1..31	
	/label="Signal peptide	
	/note="Cleavage results in partially-processed VEGF-C protein (29 kD)"	
Peptide	32..103	
	/label="Amino-terminal-peptide	
	/note="Cleavage results in fully-processed mature VEGF-C protein (21-23 kD)"	
Protein	104..227	
	/label="Mature-human-VEGF-C	
	/note="Processed vascular epithelial growth factor-C"	
Binding-site	83	
	/note="Essential for VEGFR-2 and VEGFR-3 binding"	
Active-site	131..211	
	/note="Essential for biological activity of protein"	
Binding-site	137	
	/note="Essential for VEGFR-2 and VEGFR-3 binding"	
Binding-site	156	
	/note="Essential for VEGFR-2 binding"	
Binding-site	165	
	/note="Essential for VEGFR-2 and VEGFR-3 binding"	
Peptide	228..419	
	/label="Carboxy-terminal peptide	
	/note="Cleavage results in partially-processed VEGF-C protein (29 kD)"	
WO200024412-A2.		
04-MAY-2000.		
26-OCT-1999;	99MO-US24054.	
26-OCT-1998;	98US-010587.	
(LUDW-)	LUDWIG INST CANCER RES.	
(UYHE-)	UNIV HELSINKI LICENSING LTD OY.	
(YLAH/)	YLA-HERTUULA S.	
Yla-heretuula S,	Alltalo K, Hiltunen MO, Jeltsch MM, Achen MG;	
WPI; 2000-350584/30.		
N-PSDB; AAD00339, AAD00353.		
Preventing	stenosis and restenosis in mammals using vascular	
endothelial	growth factor proteins or the nucleic acids encoding them -	
Claim 5;	Page 51-53; 61pp; English.	

CC The present amino acid sequence is the complete human prepro-vascular
CC endothelial growth factor (VEGF)-C. VEGF-C has the ability to stimulate
CC re-endothelialisation of an injured blood vessel, without significant
CC stimulation of smooth muscle cell proliferation. It can bind to and
CC stimulate VEGFR-2 (vascular endothelial growth factor receptor) and/or
CC VEGFR-3 phosphorylation in cells that express such receptors. An
CC anti-stenosis agent comprising either a VEGF-C gene or protein is
CC used in a method to reduce or prevent restenosis and stenosis of a blood
CC vessel following vascular trauma e.g., cardiovascular surgery and
CC percutaneous transluminal coronary angioplasty.

XX Sequence 419 AA:

Query Match 100.0% Score 2336; DB 21; Length 419;
Best Local Similarity 100.0%; Pred. NO. 8.9e-173;
Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MHLLGFFSVACSLAAALPPREAPAAAAAFESGLDLSDAEPDAGEATAYASKDLEOL 60
DB 1 mhlilgffsvacslaaalppreapaaaaafesglidsdaepdageatayaskdleql 60
OY 61 RSVSSVDELMVLYPEYWKMYKCOLRKGGWQHNRDQANLNSRTEETIKFAAAHYNTEILK 120
DB 61 rsvssvdelmtvlypeywkmykcqlrkqgwnreganlnsrteetikfaaahynteilk 120
OY 121 STDNEMRKTCQCPREVCIDVGEFVATNTFFKPCVSVYRGGCGNSRGLQCMNTSTSY 180
DB 121 stdnemrktqcprevcidvgefvatntffkpcvsvyrcgscnseglqcmntstsy 180
OY 181 LSKTLFEITVPLSGPKPVTISFANHTSCRCMSKLDVYRQVHSIRRSPLATLPCCOAN 240
DB 181 lsktlfeitvplsgpkpvtisfanhtscrcmskldvyrvhsirrsplatlpcqgan 240
OY 241 KTCPTNVMNNHICRCLAQEDFMFSSDAGDDSTDFGPHDTCGNKLEDETCCVCRAQIR 300
DB 241 ktcptnymnnhircclaqedfmfssdagddstdfghdpcgnkelledetccvcrraqlr 300
OY 301 PASGCPHKELDNRSCQVCYCKNKLFPSCGAGNREPDNTCCQVCKRTCPNPLNGKAC 360
DB 301 pasgcphekeldnrscqvcycknklfpsgcganreidentccqvcckrtcpnplngkac 360
OY 361 ECTESPPQCKLKGKFFHHQTCSCYRRPCTNRQKACEPGFSYSEEVCRCPVSWKRRPMS 419
DB 361 ectespqckllkgkffhhqtcscyrirpctnrqkacepgfsyseevcrvcpswkrpms 419

RESULT 9
AAB37605
ID AAB37605 standard; Protein: 419 AA.

XX AC AAB37605;
XX 27-FEB-2001 (first entry)
XX DE Human VEGF-C.
XX KW Human; gene therapy: lymphatic disorder; hereditary lymphedema; FLT4;
XX KM vascular endothelial growth factor receptor-3; VEGF-C; VEGF-D;
XX KX fms-like tyrosine kinase 4.
XX OS Homo sapiens.
XX PN CA2283470-A1.
XX PD 26-SEP-2000.
XX PF 26-SEP-1999; 99CA-2283470.
XX PR 26-MAR-1999; 99MO-US06133.
XX PR 16-AUG-1999; 99US-0375248.
XX PA (UVEPI-) UNIV PITTSBURGH.

PA (UVEH-) UNIV HELSINKI LICENSING LTD OY.
PA (LUDW-) LUDWIG INST CANCER RES.

XX Alltalo K, Ferrell RE, Finegold DN, Karkainen M;

XX WPI; 2001-007762/02.

XX N-PSDB; AAC68953.

PT Screening a human for an increased risk of developing lymphatic
PT disorder comprises assaying nucleic acid for alterations in the
PT sequences expressing vascular endothelial growth factor receptor-3
PS Disclosure; Pages 62-63; 99pp; English.

CC The present invention relates to a method for screening a human subject
CC for an increased risk of developing a lymphatic disorder e.g. hereditary
CC lymphedema. The method comprises assaying nucleic acid of a human
CC subject to determine a presence or an absence of a mutation altering the
CC sequence or expression of vascular endothelial growth factor receptor-3
CC (VEGFR-3)/fms-like tyrosine kinase 4 (FLT4) allele (see AAC68952 and
CC AAB37604) and determining an increased risk of developing lymphatic
CC disorder from presence or absence of the mutation. The presence of a
CC mutation altering the encoded amino acid sequence or expression of at
CC least 1 VEGFR-3 allele in the nucleic acid correlates with an increased
CC risk of developing a lymphatic disorder. Treatment for hereditary
CC lymphedema can be provided through the administration of vascular
CC endothelial growth factor C (VEGF-C) and vascular endothelial growth
CC factor D VEGF-D genes (via gene therapy) and proteins. The present
CC sequence is the protein sequence for VEGF-C.

XX Sequence 419 AA:

Query Match 100.0% Score 2336; DB 22; Length 419;
Best Local Similarity 100.0%; Pred. NO. 8.9e-173;
Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MHLLGFFSVACSLAAALPPREAPAAAAAFESGLDLSDAEPDAGEATAYASKDLEOL 60
DB 1 mhlilgffsvacslaaalppreapaaaaafesglidsdaepdageatayaskdleql 60
OY 61 RSVSSVDELMVLYPEYWKMYKCOLRKGGWQHNRDQANLNSRTEETIKFAAAHYNTEILK 120
DB 61 rsvssvdelmtvlypeywkmykcqlrkqgwnreganlnsrteetikfaaahynteilk 120
OY 121 STDNEMRKTCQCPREVCIDVGEFVATNTFFKPCVSVYRGGCGNSRGLQCMNTSTSY 180
DB 121 stdnemrktqcprevcidvgefvatntffkpcvsvyrcgscnseglqcmntstsy 180
OY 181 LSKTLFEITVPLSGPKPVTISFANHTSCRCMSKLDVYRQVHSIRRSPLATLPCCOAN 240
DB 181 lsktlfeitvplsgpkpvtisfanhtscrcmskldvyrvhsirrsplatlpcqgan 240
OY 241 KTCPTNVMNNHICRCLAQEDFMFSSDAGDDSTDFGPHDTCGNKLEDETCCVCRAQIR 300
DB 241 ktcptnymnnhircclaqedfmfssdagddstdfghdpcgnkelledetccvcrraqlr 300
OY 301 PASGCPHKELDNRSCQVCYCKNKLFPSCGAGNREPDNTCCQVCKRTCPNPLNGKAC 360
DB 301 pasgcphekeldnrscqvcycknklfpsgcganreidentccqvcckrtcpnplngkac 360
OY 361 ECTESPPQCKLKGKFFHHQTCSCYRRPCTNRQKACEPGFSYSEEVCRCPVSWKRRPMS 419
DB 361 ectespqckllkgkffhhqtcscyrirpctnrqkacepgfsyseevcrvcpswkrpms 419

RESULT 10

XX AAM13833
XX ID AAM13833 standard; Protein: 419 AA.
XX AC AAM13833;
XX DT 05-JUN-1997 (first entry)

```

XX Human vascular endothelial growth factor-related protein VRP.
DE
XX
XX Vascular endothelial growth factor-related protein; VRP; VEGF;
KM receptor protein tyrosine kinase; Flt4; signal transduction;
KM wound healing; vulnerrary; rheumatoid arthritis; Kaposi's sarcoma;
KM therapy; diagnosis; angiogenesis; monoclonal antibody.
XX
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH Peptide 1..20
FT /label= Sig_peptide
FT Protein 20..419
FT /label= Mat_protein
FT Misc-difference 114
FT /note= "deduced residue from nucleotide sequence
FT is tyrosine"
XX
XX WO9709427-A1.
XX
XX 13-MAR-1997.
XX
XX 30-AUG-1996; 96WO-US14075.
XX
XX 08-SEP-1995; 95US-0003491.
XX
XX (GETH ) GENENTECH INC.
XX
XX Lee J, Wood W;
XX
XX WPI: 1997-192902/17.
XX
XX N-PSDB: AAT59929.
XX
XX Human protein similar to vascular endothelial growth factor - used
XX to treat e.g. wounds, tumours, rheumatoid arthritis, Kaposi's
XX sarcoma etc.
XX
XX Claim 6: Fig 1A-D: 68pp; English.
XX
XX A human vascular endothelial growth factor (VEGF)-related protein
XX (VRP) (AAW13833) has been identified that binds to, and stimulates
XX the phosphorylation of, the receptor tyrosine kinase Flt4. It is
XX postulated to be a third member of the VEGF protein family. Its
XX amino acid sequence was deduced from a cDNA clone (AAT59929) obtd.
XX from a glioma G61 library. Recombinant VRP can be produced in
XX transformed host cells and used: to promote growth of vascular and
XX lymph endothelial cells; to stimulate phosphorylation of the
XX tyrosine kinase domain of a Flt4 receptor; as a diagnostic; as an
XX additive to cell cultures; to screen for (ant)agonists; and to
XX raise monoclonal antibodies used to treat conditions associated
XX with excessive neovascularisation or vascular permeability. VRP
XX may make it possible to avoid coronary by-pass surgery by
XX stimulating growth of the collateral circulation.
XX
XX Sequence 419 AA:
SQ

```

```

Query Match          99.6%; Score 2327; DB 18; Length 419;
Best Local Similarity 99.8%; Pred. No. 4.4e-172;
Matches 418; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 181 LSKTLEITWPLSQGPPTISFANHTSCRMKSLDVRQVHSIIIRSLPATLPQCOAN 240
DB 181 LSKTLEITWPLSQGPPTISFANHTSCRMKSLDVRQVHSIIIRSLPATLPQCOAN 240
OY 241 KTCPTVYMNHNHICRLAODEFMFSSDAGSDTDFGHDICGPNKELDEFTCCVCBAGLR 300
DB 241 KTCPTVYMNHNHICRLAODEFMFSSDAGSDTDFGHDICGPNKELDEFTCCVCBAGLR 300
OY 301 PASCGPHKELDRNSCCVCVCKNKLFPSCCGANRPFEDNTCCVCCKRCPNPNPKCKAC 360
DB 301 PASCGPHKELDRNSCCVCVCKNKLFPSCCGANRPFEDNTCCVCCKRCPNPNPKCKAC 360
OY 361 ECTESPCKLKGKRFHQTSCCYRRPCTNPOKACEPGFSEYCVRCVPSYWKRRQMS 419
DB 361 ECTESPCKLKGKRFHQTSCCYRRPCTNPOKACEPGFSEYCVRCVPSYWKRRQMS 419

RESULT 11
ID AAY30518 standard; Protein; 419 AA.
XX
XX AAY30518;
XX
XX 16-NOV-1999 (first entry)
XX
XX Vascular endothelial growth factor-2 (VEGF-2).
XX
XX Human vascular endothelial growth factor-2; VEGF-2;
XX Vascular endothelial cell growth; endothelial cell migration;
XX angiogenesis; blood pressure; blood flow; immune system disorder;
XX immune cell; cancer; autoimmune disorder; blood protein disorder;
XX ataxia telangiectasia; common variable immunodeficiency;
XX DisGeorge syndrome; HIV infection; HTLV-BIV infection;
XX leukocyte adhesion deficiency syndrome; lymphopenia;
XX phagocyte bactericidal dysfunction; severe combined immunodeficiency;
XX Wiskott-Aldrich disorder; anemia; thrombocytopenia; hemoglobinuria;
XX allergy; asthma; allergic asthma.
XX
XX Homo sapiens.
XX
XX WO9946364-A1.
XX
XX 16-SEP-1999.
XX
XX 10-MAR-1999; 99WO-US05021.
XX
XX 13-MAR-1998; 98US-0042105.
XX
XX 30-JUN-1998; 98US-0107997.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Cao L, Hu J;
XX
XX WPI: 1999-551399/46.
XX
XX N-PSDB: AAZ10523.
XX
XX New human vascular endothelial growth factor-2, used for treating, e.g.
XX immune disorders and cancers.
XX
XX Claim 12; Fig 1A-E; 222pp; English.
XX
XX The present sequence represents vascular endothelial growth factor-2
XX (VEGF-2). The VEGF-2 polypeptides have activities similar to VEGF. The
XX VEGF-2 polypeptides stimulate the growth of vascular endothelial cells,
XX stimulate endothelial cell migration, stimulate angiogenesis, decrease
XX blood pressure, and increase blood flow. The polynucleotides and
XX polypeptides can be used for preventing, treating or ameliorating a
XX medical condition. The VEGF-2 polypeptides or polynucleotides may be
XX useful in treating deficiencies or disorders of the immune system, by
XX activating or inhibiting the proliferation, differentiation, or
XX mobilization (chemotaxis) of immune cells. The etiology of these immune
XX deficiencies or disorders may be genetic, somatic, such as cancer or

```

CC some autoimmune disorders, acquired (e.g. by chemotherapy or toxins), or
CC infections. Examples of immunologic deficiency syndromes include blood
CC protein disorders, ataxia telangiectasia, common variable
CC immunodeficiency, DiGeorge syndrome, HIV infection, HTLV-BLV infection,
CC leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte
CC bactericidal dysfunction, severe combined immunodeficiency (SCIDs),
CC Minkowski-Richthofen disorder, anemia, thrombocytopenia, or hemoglobinuria.
CC They can also be used to modulate emostatic or thrombolytic activity.
CC Similarly allergic reactions and conditions such as asthma (particularly
CC allergic asthma) or other respiratory problems, may also be treated.

Sequence 419 AA;

Query Match	99.68;	Score 2326;	DB 20;	Length 419;
Best Local Similarity	99.58;	Pred. No. 5.3e-172;		
Matches 417; Conservative	1;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	MLHGFVSACSLIAAALLBPBRAPAAAAFESGIDLSDAEPDAGATVYASDLDEOL	60
Db	1	mhs.igffsvaasllaaallbpptreapaaaafesgldlsdaepdaagaaayaskdxldeql	60
QY	61	RSVSVDELMTVLYPEYWKWKCKOLRRGGGOMHNEQANLSRTEETLTKFAAHYNTIELK	120
Db	61	rsvsvdelmtvlypeywkmykcqlrlyrgygmhneqanlsrteetlkfkaahyntelk	120
QY	121	SIDNEWRKTQCMPEVCDIDVGEFVGATNTPEKPCSVYRCGGCCNSEGLQCMNTSTSY	180
Db	121	sidnewrktqcmpevcidvgefgvatntfpekpccsvyrcggccnsegldcmntstsy	180
QY	181	LSKTLFEITVPLSGRPKPVITSPFNHNSCHMSKSLDYRQVHSTRSLPATLPCQQAAN	240
Db	181	lskltfelitvplsgrpkrpvlistanhschmskldyryvhsllrrslpatlpcqqaan	240
QY	241	KTCPTNTMMNNHILCRCAOEDFMFSSDAGSDSTGFIHDIGPKMKELDEEPCOCVCRAGR	300
Db	241	ktcptntmmnnhiltcrcaoledfmfssdagdsdtdgfhdicgpkmkeldeepcoccvcragr	300
QY	301	PASGPKHELDNRNSCOCVCKNKLFPSSCGGNREDENTCOVCRTCPRNQPLNPVKCAC	360
Db	301	pascgpkhieldnrnsccvcvcknklfpsgcggnreidentcgvcvckrtcpnqplnpvkcac	360
QY	361	ECTSSPOKLLKGGKFNHJQSCSTRRPCTNRQKACERGESYSEEVCRQCVSYWKKRQMS	419
Db	361	ectsspckllkgykfnhjqscstrrpctnrrqkacergefyseevcrqcvsvywkkrrpms	419

RESULT 12
 AAY22320

AC MAY22320;

DT 22-SEP-1999 (first entry)

DE Full length human VEGF2 protein sequence.

KM VEGF2; vascular endothelial growth factor 2; angiogenesis; bone damage;
KM endothelial cell proliferation; tissue damage; therapy.

OS Homo sapiens.

PN US5932540-A

PD 03-AUG-1999

PF 24-DEC-1997; 97US-0999811.

PR 24-DEC-1997; 97US-0999811.

PR	06-JUN-1995:	95US-0465968.
FN	06-MAR-1994,	94US-0207220.

AA
PA (HUMA-) HUMAN GENOME SCI INC

XX Cao L, Hu J, Rosen CA;
PI WPI; 1999-443606/37.
XX DR N-PSDB; AAX84837.
DR

PT Vascular endothelial growth factor 2 for wound healing and vascular
PT repair

PS Claim 1; Fig 1; 49pp; English.

CC This sequence is the vascular endothelial growth factor 2 (VEGF2),
CC of the invention. The isolated polypeptide is useful for stimulating
CC angiogenesis, by promoting the proliferation of endothelial cells, for
CC the treatment of a wound, or for the treatment of tissue or bone damage

Sequence 419 AA;

Query Match	99.6%;	Score 2326;	DB 20;	Length 419;
Best Local Similarity	99.5%;	Pred. No. 5.3e-172;		
Matches 417;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	MHGGFESVACSLIAAALPEGPREPAPAAAAREGIDISDEBPAGEXTAVASMDLEOL	60
Db	1	mhs1gffsvacs1laaal1pgpreapaaaaa1esg1d1staepdageatayaskd1eeq1	60
QY	61	RSVSSVDELMLVLYPEVWKMYCOLRKGGWOMHNEQALNSTRTEETIFAAAHNTIELK	120
Db	61	rsvsvsdelmtvly1pevykmykcol1rkggwomhneqalnstrteetlk1faahyntelk1	120
QY	121	SIDNEWRKTCQMPREVCIDVGKEFGVATNTFFKPPCVSVYRCGCCNSEGLQMNSTSY	180
Db	121	sidnewrktcqmprevcidvgk1efgvatntffkppcvsvyrcgcsnseglqmntstsy1	180
QY	181	LSKLFLETTVPLSOGPKPVYTSFSAFNHNSCRMSKLDYRYOYHSHIRSLPATLEPQOQAN	240
Db	181	lsk1lfeltvpl1sgpkrpvt1sfanfhnscrmkskldyryghsh1rrslpat1pqqaan1	240
QY	241	KTCPTNYMNMNHHICRCLAODEFMFSADGSDSTGGFHICGPNKLEDEETOCYCRAGLR	300
Db	241	ktcptnynmnh1crrclaogedfmf1ssdagds1tdg1fnd1cgnk1eldeet1cvc1raglr1	300
QY	301	PASGCPHKEIDRNSCQCCKNKLPPSOGANREFEDNTCQVCCKRTCPRNDPLNPGCAC	360
Db	301	pascgphkeidrnscqvccknkl1fsgcganrefednt1cgvckrtcp1rnp1npgcac1	360
QY	361	ECTSSPQCKLLKGGKFFHHQTCSCRRRECTNQNQKACBEGFSISEVYCAVSVYKRRPWS	419
Db	361	ectsspqckllk1ggkffhhqtcsc1rrrectnqnqkac1begfs1sevycavsvy1krrpws1	419

RESULT	13
AAV97144	
ID	AAV97144 standard; Protein; 419 AA

AC AAY97144;

DT 22-DEC-2000 (first entry)

DE Vascular endothelial growth factor-2 (VEGF-2).

KM Vascular endothelial growth factor 2; VEGF-2; retina; angiogenesis
KM treatment; injury; degeneration; photoreceptors; eye;
KM angioid streaks; retinitis; pigmentosa; human;
KM age-related macular degeneration; diabetic retinopathy.

OS Homo sapiens.

PN WO2000045835-A1

PD 10-AUG-2000


```

PF 07-FEB-2000; 2000MO-US03047.
XX
XX 08-FEB-1999; 99US-0119179.
PR 12-FEB-1999; 99US-0119926.
PR 03-JUN-1999; 99US-0137796.
PR 22-DEC-1999; 99US-0171505.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Alderson R, Melder R, Roschke V, Ruben SM;
PI WPI; 2000-532862/48.
DR N-PSDB; AAA52080.
XX
XX Treating injury or degeneration of photoreceptors comprises
PT administering to a subject vascular endothelial growth factor 2
PT (VEGF-2)
XX
XX Claim 31: Fig 1a-e; 252pp; English.
XX
XX Administration of vascular endothelial growth factor 2 (VEGF-2)
CC to a patient can be used for treating injury or degeneration of
CC photoreceptors associated with e.g. angiod streaks, retinitis
CC pigmentosa, age-related macular degeneration, diabetic retinopathy,
CC etc. VEGF-2 promotes angiogenesis, the formation of new blood
CC vessels in the retina.
XX
XX Sequence 419 AA:
SQ
Query Match 99.6%; Score 2326; DB 21; Length 419;
Best Local Similarity 99.5%; Pred. No. 5.3e-172;
Matches 417; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 MHLGFEFVACSLAAALLPGPREAPAAAAAFESGLDSDAPDAGEATAYASKDLEQL 60
DB 1 mnslglfsvacsliaaallppreapaaataesgidlsdaepdageataaskdleeql 60
OY 61 RSVSSVDELMTVLYPEYWKMYKCOLRKGGWQHNRDOANLNRTEETIKFAAHYTEILK 120
DB 61 rsvssvdelmtvlypeywkmykcolrkggwqhnrqanlnsteeetikfaaahynteilk 120
OY 121 STDNEMRKTCMPREVCIDVKEFGVATNTFFKPCVSVYRCGGCNSBGLOCMNTSTSY 180
DB 121 stdnemrktcmprevcidvgkefgvatntffkpcvsvyrcggcnsbglocmntstsy 180
OY 181 LSKTLEITVPLSQGPKPYTISFANTSCRCMSKLDVYRQVHSITRRSLPATLPQCOAAN 240
DB 181 lsktleitvplsqgpkpytisfantscrcmskldvyrvhsitrrslpatlpqcoaan 240
OY 241 KTCPTVMNNNNHICRCLAODEFMFSSDAGDSTDFHDI CGPNKELDETCQCVCRAGLR 300
DB 241 ktcpvmnnnnhrcrlaodefmfssdagdstdfhdi cgpnlkdeetccqcvcraglr 300
OY 301 PASCGHAKELDRNSCCVCCKNLFPSCGAGNREPDENTOCQVCKRTCPRNQPLNPGKAC 360
DB 301 pascgahakeldrnsccvccknlfpscgagndepdentocqvcckrtcpnqlnpkacac 360
OY 361 ECESPOKCLLGGKRRHHOTGSCYRRPCTNNROKACBPGFSYSREVCRCVPSWKKRQMS 419
DB 361 ecespo kcll gkrrhh otgsc yrrpctnn rokac bpgfsys revcrcvpswkk r qms 419
RESULT 14
AA97570
ID AA97570 standard; Protein; 419 AA.
XX
XX AA97570;
AC AA97570;
XX
XX 05-APR-2001 (first entry)
XX
XX Human VEGF-B protein sequence.
DE
XX

```

```

KM Human; angiogenic protein; wound healing; vascular tissue repair;
KM peripheral arterial disease; critical limb ischaemia; coronary disease;
KM angiogenesis; tumour; inflammation; diabetic retinopathy; psoriasis;
KM rheumatoid arthritis; autoimmune disease; allergy; cancer; therapy;
KM infectious disease; neurodegeneration;
KM vascular endothelial growth factor-B; VEGF-B.
XX
XX Homo sapiens.
XX
XX WO200075163-A1.
XX
XX 14-DEC-2000.
XX
XX 01-JUN-2000; 2000MO-US14925.
XX
XX 03-JUN-1999; 99US-0137796.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Hu J, Cao L;
PI WPI; 2001-071057/08.
DR N-PSDB; AAA91004.
XX
XX New nucleic acid encoding angiogenic proteins, useful e.g. for
PT promoting healing of wounds and treating peripheral arterial disease,
PT critical limb ischaemia or coronary disease -
XX
XX Claim 11: Fig 1; 244pp; English.
XX
XX This sequence is vascular endothelial growth factor-B (VEGF-B),
XX which is an angiogenic protein of the invention. The angiogenic proteins
XX and the DNA sequences encoding them, are used to prevent, treat or
XX ameliorate disease and to detect diseases, or susceptibility, by
XX detecting mutations or the presence or amount of angiogenic protein
XX expression. Particularly they are used to stimulate wound healing,
XX growth of damaged bone and tissue, and for repair of vascular tissue,
XX especially peripheral arterial disease, critical limb ischaemia or
XX coronary disease. Antagonists of the sequences are used to inhibit
XX angiogenesis in tumours and to treat inflammation (where associated with
XX increased vascular permeability), diabetic retinopathy, rheumatoid
XX arthritis or psoriasis. Agonists are also useful for stimulating
XX (lymph)angiogenesis. The proteins are also used to identify specific
XX binding agents (potential therapeutic agents) and to raise antibodies.
XX The antibodies are useful as therapeutic (ant)agonists; for detection,
XX purification and targeting of proteins for in vivo or in vitro diagnosis
XX (including imaging) or for therapy (including when linked to e.g. a label
XX or cytotoxin); and for immunotyping of cells, e.g. for detecting minimal
XX residual disease or haematopoietic progenitor/stem cells. It is also
XX contemplated that the sequences might be useful for treating a very wide
XX range of other disorders, e.g. autoimmune diseases; allergy; cancer;
XX infectious diseases (viral, bacterial, fungal or parasitic);
XX neurodegeneration, also as chemotactic agents or for stimulating
XX regeneration of the nervous system etc.
XX
XX Sequence 419 AA:
SQ
Query Match 99.6%; Score 2326; DB 22; Length 419;
Best Local Similarity 99.5%; Pred. No. 5.3e-172;
Matches 417; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 MHLGFEFVACSLAAALLPGPREAPAAAAAFESGLDSDAPDAGEATAYASKDLEQL 60
DB 1 mnslglfsvacsliaaallppreapaaataesgidlsdaepdageataaskdleeql 60
OY 61 RSVSSVDELMTVLYPEYWKMYKCOLRKGGWQHNRDOANLNRTEETIKFAAHYTEILK 120
DB 61 rsvssvdelmtvlypeywkmykcolrkggwqhnrqanlnsteeetikfaaahynteilk 120
OY 121 STDNEMRKTCMPREVCIDVKEFGVATNTFFKPCVSVYRCGGCNSBGLOCMNTSTSY 180
DB 121 stdnemrktcmprevcidvgkefgvatntffkpcvsvyrcggcnsbglocmntstsy 180

```



```

OY 181 LSKTLFEITVPLSGPKPVITISFANHSCRCMSKIDVYRQVHSIIRSLPATLPCCQAN 240
    |||||||
DB 181 lsktlfeitvplsgpkpvtisfanhscrcmsklidyryqvhsilrrslpatlpccqaan 240
OY 241 KTCPTNYMNMNHHICRCLAOEDFMFSSDAGDDSTDFPHDICGPNKELDEETCCVCRCAGLR 300
    |||||||
DB 241 ktcptnymnmnhicrcclaoedfmfssdagddstdfphdicgpnkeldeetccvcrcaglr 300
OY 301 PASCGRKELDRNSQCVCCKNKLFPSCGAGNREFDENTCQCCKRTCPNRPPLNGKAC 360
    |||||||
DB 301 pascgrkeldrnscqvccknklfpscgaganeidentcqcckrtcpnrgplngkcac 360
OY 361 ECTESPOKCLTKGKKFHHQTCSCYRRCPTNRKACEPGEFSYSEEVCRCPVSWKRPQMS 419
    |||||||
DB 361 ectespgkcllkgkfhqtcscyrpctnrkacepgfsyseevcrvpswkrpqms 419

RESULT 15
AAM75751
ID AAM75751 standard; Protein: 419 AA.
XX
AC AAM75751;
XX
DT 14-DEC-1998 (first entry)
XX
DE Vascular endothelial growth factor C protein analogue.
XX
KW Fil4; vascular endothelial growth factor C; vascular endothelial cell;
    lymphatic endothelial cell; myelopoiesis; angiogenesis; inflammation;
    lymphangiogenesis; oedema; elephantiasis; Milroy's disease.
XX
OS Synthetic.
XX
FH Key location/qualifiers
FT Modified-site 136
FT /note= "Xaa can be anything other than cysteine, or
    can be nothing"
XX
PN W09833917-A1.
XX
PD 06-AUG-1998.
XX
PE 02-FEB-1998; 98MO-US01973.
XX
PR 05-FEB-1997; 97US-0795430.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
    (UYHE-) UNIV HELSINKI LICENSING LTD.
XX
PI Alitalo K, Joukov V;
XX
DR WPI: 1998-437470/37.
XX
PT New isolated vascular endothelial growth factor polypeptide(s) -
    used to develop products for treating, e.g. cancers, inflammation,
    oedema, granulocytopenia or for wound healing or tissue
    transplantation
XX
PS Example 35; Page 143-145; 177bp; English.
XX
CC The vascular endothelial growth factor C (VEGF-C) polypeptides have
    activities affecting growth and migration of vascular endothelial cells,
    promoting growth of lymphatic endothelial cells and lymphatic vessels,
    increasing vascular permeability, and affecting myelopoiesis. The
    products can be used for stimulating angiogenesis, for inhibiting
    angiogenesis, for stimulating lymphangiogenesis, treatment or prevention
    of inflammation, oedema, elephantiasis, or Milroy's disease. They can
    also be used to modulate myelopoiesis, e.g. treating granulocytopenia.
    They can also be used for modulating the growth of endothelial cells.
    They can also be used to stimulate lymphocyte production and maturation,
    and to promote or inhibit trafficking of leucocytes between tissues and
    lymphatic vessels or to affect migration in and out of the thymus.

```

```

XX SQ Sequence 419 AA;
Query Match 99.5%; Score 2325; DB 19; Length 419;
Best Local Similarity 99.8%; Pred. No. 6.3e-172;
Matches 418; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MHLGFFSVACSLAAALPPGREAFAAAAFEGSLDSDAEPDAGETAYASKDLEQL 60
    |||||||
DB 1 mlllgffsvacslaaalppgreaaafesglldisdaepdagetaaskdleeql 60
OY 61 RSVSSVDELMETVLYPEYKMKYCOLRKGMQHNREQANLNSRTEETIRFAAHYNTETILK 120
    |||||||
DB 61 rsvssvdelmetvlypeykmkycolrkqgmhnrqanlnsrteetirfaahyntetilk 120
OY 121 SIDNEMRKTCQMPREVCIDVCKEFGVATNTPFKPPCVSVYRCGCCNSEGLQCMNTSTSY 180
    |||||||
DB 121 sidnemrktqcmprcvcidvckefgvatntpfkppcvsvyrcgccnseglqcmntstsy 180
OY 181 LSKTLFEITVPLSGPKPVITISFANHSCRCMSKIDVYRQVHSIIRSLPATLPCCQAN 240
    |||||||
DB 181 lsktlfeitvplsgpkpvtisfanhscrcmsklidyryqvhsilrrslpatlpccqaan 240
OY 241 KTCPTNYMNMNHHICRCLAOEDFMFSSDAGDDSTDFPHDICGPNKELDEETCCVCRCAGLR 300
    |||||||
DB 241 ktcptnymnmnhicrcclaoedfmfssdagddstdfphdicgpnkeldeetccvcrcaglr 300
OY 301 PASCGRKELDRNSQCVCCKNKLFPSCGAGNREFDENTCQCCKRTCPNRPPLNGKAC 360
    |||||||
DB 301 pascgrkeldrnscqvccknklfpscgaganeidentcqcckrtcpnrgplngkcac 360
OY 361 ECTESPOKCLTKGKKFHHQTCSCYRRCPTNRKACEPGEFSYSEEVCRCPVSWKRPQMS 419
    |||||||
DB 361 ectespgkcllkgkfhqtcscyrpctnrkacepgfsyseevcrvpswkrpqms 419

```

Search completed: October 17, 2001, 14:45:59
Job time: 125 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:43:54 ; Search time 62.93 Seconds
(without alignments)
137.094 Million cell updates/sec

Title: US-09-427-657-2

Perfect score: 2336
Sequence: 1 MHLGFRVACSLIAALLP.....SYSEVCRVPSYWKRPQMS 419

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued Patents-AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2336	100.0	419	4	US-08-795-430-8
2	2336	100.0	419	4	US-08-510-133A-35
3	2326	99.6	419	2	US-08-999-811-2
4	2326	99.6	419	3	US-09-042-105-2
5	2326	99.6	419	3	US-09-042-105-18
6	2320	99.3	419	5	PCT-US96-09001-2
7	2048	87.7	415	4	US-08-795-430-11
8	1999	85.6	350	4	US-08-510-133A-33
9	1999	85.6	350	4	US-08-585-895-33
10	1995	85.4	350	2	US-08-999-811-4
11	1995	85.4	350	2	US-08-824-996-2
12	1995	85.4	350	3	US-09-042-105-4
13	1804.5	77.2	418	4	US-08-795-430-13
14	704.5	30.2	325	4	US-08-915-795-3
15	704.5	30.2	354	4	US-08-915-795-5
16	703.5	30.1	358	4	US-08-915-795-8
17	667.5	28.6	321	4	US-08-915-795-9
18	238.5	10.2	232	2	US-08-999-811-7
19	238.5	10.2	232	3	US-08-807-992B-4
20	238.5	10.2	232	3	US-09-042-105-7
21	238	10.2	231	2	PCT-US96-09001-10
22	235.5	10.1	232	5	US-08-824-996-9
23	228.5	9.8	214	6	5240848-11
24	226	9.7	215	3	US-08-807-992B-3
25	226	9.7	215	4	US-08-586-039B-49
26	226	9.7	215	6	5240848-7
27	222	9.5	191	3	US-08-567-200A-2

28	222	9.5	191	3	US-08-807-992B-2	Sequence 2, Appl1
29	222	9.5	191	3	US-08-691-794-2	Sequence 2, Appl1
30	222	9.5	191	4	US-08-795-430-56	Sequence 56, Appl1
31	222	9.5	191	6	5332671-4	Patent No. 5332671
32	221.5	9.5	165	6	5219739-22	Patent No. 5219739
33	217.5	9.3	165	6	5194596-18	Patent No. 5194596
34	217.5	9.3	165	6	5219739-19	Patent No. 5219739
35	210	9.0	189	1	US-08-469-427A-15	Sequence 15, Appl1
36	206.5	8.8	190	4	US-08-586-039B-31	Sequence 31, Appl1
37	206.5	8.8	214	4	US-08-586-039B-35	Sequence 35, Appl1
38	205.5	8.8	190	6	5332671-3	Patent No. 5332671
39	202.5	8.7	164	6	5194596-17	Patent No. 5194596
40	202.5	8.7	164	6	5219739-17	Patent No. 5219739
41	202.5	8.7	164	6	5219739-18	Patent No. 5219739
42	198.5	8.5	190	2	US-08-565-063C-20	Sequence 20, Appl1
43	192	8.2	147	3	US-08-807-992B-1	Sequence 1, Appl1
44	190.5	8.2	145	3	US-08-784-551C-2	Sequence 2, Appl1
45	187.5	8.0	121	6	5194596-19	Patent No. 5194596

ALIGNMENTS

RESULT 1
US-08-795-430-8
Sequence 8, Application US/08795430
Patent No. 6130071
GENERAL INFORMATION:
APPLICANT: Allitalo, Karl
APPLICANT: Joukov, Vladimir
TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
TITLE OF INVENTION: Protein and Gene, Mutants Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/795,430
APPLICATION NUMBER: US/08/795,430
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FI96/00427
FILING DATE: 01-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/671,573
FILING DATE: 28-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/601,132
FILING DATE: 14-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/585,895
FILING DATE: 12-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/510,133
FILING DATE: 01-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,011
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Cass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28967/33691
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448
 TELEEX: 25-3856
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 419 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-795-430-8

Query Match 100.0%; Score 2336; DB 4; Length 419;
 Best Local Similarity 100.0%; Pred. No. 1.7e-202;
 Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHLGFFSVACSLAAALPGPREAPAAAAAFESGIDLSDAEPDAGEATAYASKDLEQL 60
 DB 1 MHLGFFSVACSLAAALPGPREAPAAAAAFESGIDLSDAEPDAGEATAYASKDLEQL 60
 QY 61 RSVSSVDELMTVLYPEYWMYKCOLRKGGMOHNRQOANINSRTEETIKFAAAHYNTIELK 120
 DB 61 RSVSSVDELMTVLYPEYWMYKCOLRKGGMOHNRQOANINSRTEETIKFAAAHYNTIELK 120
 QY 121 SIDNEMRKTCQMPREVCIQVGEFVATNTFFKPCVSVYRCGGCCNSGLQCMNTSTSY 180
 DB 121 SIDNEMRKTCQMPREVCIQVGEFVATNTFFKPCVSVYRCGGCCNSGLQCMNTSTSY 180
 QY 181 LSKTLEITVPLSQGPKPTISFANTSCRCMSKLDVYQVHSIIRSLPATLPQOQAN 240
 DB 181 LSKTLEITVPLSQGPKPTISFANTSCRCMSKLDVYQVHSIIRSLPATLPQOQAN 240
 QY 241 KTCPTVMNNNHICRLAEDFMFSSDAGDDSTDFHICGPKNELDEETCCVCAGLR 300
 DB 241 KTCPTVMNNNHICRLAEDFMFSSDAGDDSTDFHICGPKNELDEETCCVCAGLR 300
 QY 301 PASGCPHKLDRNSCCQVCKNKLFPSCGANNREFDENTCCQVCKRTCPNPNPLNPGKAC 360
 DB 301 PASGCPHKLDRNSCCQVCKNKLFPSCGANNREFDENTCCQVCKRTCPNPNPLNPGKAC 360
 QY 361 ECTESPQKCLLGGKFFHQTSCYRRPCTNRQACGPGFSYSEVRCVPSYWKRPQMS 419
 DB 361 ECTESPQKCLLGGKFFHQTSCYRRPCTNRQACGPGFSYSEVRCVPSYWKRPQMS 419

RESULT 2

US-08-510-133A-35
 Sequence 35, Application US/08510133A

Patent No. 6221839
 GENERAL INFORMATION:

APPLICANT: Alltalo, Karl
 Joukov, Vladomir
 TITLE OF INVENTION: Receptor Ligand
 NUMBER OF SEQUENCES: 35
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 6300 Sears Tower, 233 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/510,133A

FILING DATE: 01-Aug-1995

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Gass, David A.

REGISTRATION NUMBER: 38,153

REFERENCE/DOCKET NUMBER: 28113/32863

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/474-6300
 TELEFAX: 312/474-0448
 TELEEX: 25-3856
 INFORMATION FOR SEQ ID NO: 35:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 419 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 35:
 US-08-510-133A-35

Query Match 100.0%; Score 2336; DB 4; Length 419;
 Best Local Similarity 100.0%; Pred. No. 1.7e-202;
 Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHLGFFSVACSLAAALPGPREAPAAAAAFESGIDLSDAEPDAGEATAYASKDLEQL 60
 DB 1 MHLGFFSVACSLAAALPGPREAPAAAAAFESGIDLSDAEPDAGEATAYASKDLEQL 60
 QY 61 RSVSSVDELMTVLYPEYWMYKCOLRKGGMOHNRQOANINSRTEETIKFAAAHYNTIELK 120
 DB 61 RSVSSVDELMTVLYPEYWMYKCOLRKGGMOHNRQOANINSRTEETIKFAAAHYNTIELK 120
 QY 121 SIDNEMRKTCQMPREVCIQVGEFVATNTFFKPCVSVYRCGGCCNSGLQCMNTSTSY 180
 DB 121 SIDNEMRKTCQMPREVCIQVGEFVATNTFFKPCVSVYRCGGCCNSGLQCMNTSTSY 180
 QY 181 LSKTLEITVPLSQGPKPTISFANTSCRCMSKLDVYQVHSIIRSLPATLPQOQAN 240
 DB 181 LSKTLEITVPLSQGPKPTISFANTSCRCMSKLDVYQVHSIIRSLPATLPQOQAN 240
 QY 241 KTCPTVMNNNHICRLAEDFMFSSDAGDDSTDFHICGPKNELDEETCCVCAGLR 300
 DB 241 KTCPTVMNNNHICRLAEDFMFSSDAGDDSTDFHICGPKNELDEETCCVCAGLR 300
 QY 301 PASGCPHKLDRNSCCQVCKNKLFPSCGANNREFDENTCCQVCKRTCPNPNPLNPGKAC 360
 DB 301 PASGCPHKLDRNSCCQVCKNKLFPSCGANNREFDENTCCQVCKRTCPNPNPLNPGKAC 360
 QY 361 ECTESPQKCLLGGKFFHQTSCYRRPCTNRQACGPGFSYSEVRCVPSYWKRPQMS 419
 DB 361 ECTESPQKCLLGGKFFHQTSCYRRPCTNRQACGPGFSYSEVRCVPSYWKRPQMS 419

RESULT 3

US-08-999-811-2
 Sequence 2, Application US/08999811

Patent No. 5932540
 GENERAL INFORMATION:

APPLICANT: HU, JING-SHAN
 APPLICANT: ROSEN, CRAIG A.

TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX

STREET: 1100 NEW YORK AVENUE

CITY: WASHINGTON

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/999,811

FILING DATE: HEREWITH

CLASSIFICATION:

```

PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550
FILING DATE: 8-MAR-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: MARKOWICZ, KAREN R.
REGISTRATION NUMBER: 36,351
REFERENCE/DOCKET NUMBER: 1488.1000004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-999-811-2

```

```

Query Match          99.6%; Score 2326; DB 2; Length 419;
Best Local Similarity 99.5%; Pred. No. 1.3e-201;
Matches 417; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MHLLGFSSVACSLAALLPGPREAPAAAAAFESGLDSDAEPDAGEATAYASKDLEOL 60
DB 1 MHLLGFSSVACSLAALLPGPREAPAAAAAFESGLDSDAEPDAGEATAYASKDLEOL 60
QY 61 RSVSSVDELMTVLYPEYWMYKCOLRKGGWQHNRQANLNSTREETIKFAAHYNTIELK 120
DB 61 RSVSSVDELMTVLYPEYWMYKCOLRKGGWQHNRQANLNSTREETIKFAAHYNTIELK 120
QY 121 STDNEMRKTCQCPREVCIDVGEKFGVATNTFFKPCVSYRCGGCNSGLQCMNTSTSY 180
DB 121 STDNEMRKTCQCPREVCIDVGEKFGVATNTFFKPCVSYRCGGCNSGLQCMNTSTSY 180
QY 181 LSKTLFEITVPLSOGPKPVTISFANHTSCRCMSKLDVYRQVHSIIRSLPATLPQOQAN 240
DB 181 LSKTLFEITVPLSOGPKPVTISFANHTSCRCMSKLDVYRQVHSIIRSLPATLPQOQAN 240
QY 241 KTCPTNYMNNHICRCLAODEFMFSSDAGDSTDFGHDICGPNKELDEETCCQVCRAGLR 300
DB 241 KTCPTNYMNNHICRCLAODEFMFSSDAGDSTDFGHDICGPNKELDEETCCQVCRAGLR 300
QY 301 PASGCPHRELDRNSQCVCCKNKLFPSCGANEPEDENTCCQCKRTCPRNOPLNPGKAC 360
DB 301 PASGCPHRELDRNSQCVCCKNKLFPSCGANEPEDENTCCQCKRTCPRNOPLNPGKAC 360
QY 361 ECTESPQCKLLGKKFHHOTCSYRRPCTNRQKACEPGFSYSEEVCRVPSYWRPQMS 419
DB 361 ECTESPQCKLLGKKFHHOTCSYRRPCTNRQKACEPGFSYSEEVCRVPSYWRPQMS 419

```

RESULT 4

```

US-09-042-105-2
Sequence 2, Application US/09042105
Patent No. 6040157
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESS: STERN, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:

```

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,105
FILING DATE: HEREMITH
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550
FILING DATE: 8-MAR-1994
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: TO BE ASSIGNED
FILING DATE: 24-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: ERIC K. STEFFE
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2540
TELEFAX: (202)371-2600
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-042-105-2

```

```

Query Match          99.6%; Score 2326; DB 3; Length 419;
Best Local Similarity 99.5%; Pred. No. 1.3e-201;
Matches 417; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MHLLGFSSVACSLAALLPGPREAPAAAAAFESGLDSDAEPDAGEATAYASKDLEOL 60
DB 1 MHLLGFSSVACSLAALLPGPREAPAAAAAFESGLDSDAEPDAGEATAYASKDLEOL 60
QY 61 RSVSSVDELMTVLYPEYWMYKCOLRKGGWQHNRQANLNSTREETIKFAAHYNTIELK 120
DB 61 RSVSSVDELMTVLYPEYWMYKCOLRKGGWQHNRQANLNSTREETIKFAAHYNTIELK 120
QY 121 STDNEMRKTCQCPREVCIDVGEKFGVATNTFFKPCVSYRCGGCNSGLQCMNTSTSY 180
DB 121 STDNEMRKTCQCPREVCIDVGEKFGVATNTFFKPCVSYRCGGCNSGLQCMNTSTSY 180
QY 181 LSKTLFEITVPLSOGPKPVTISFANHTSCRCMSKLDVYRQVHSIIRSLPATLPQOQAN 240
DB 181 LSKTLFEITVPLSOGPKPVTISFANHTSCRCMSKLDVYRQVHSIIRSLPATLPQOQAN 240
QY 241 KTCPTNYMNNHICRCLAODEFMFSSDAGDSTDFGHDICGPNKELDEETCCQVCRAGLR 300
DB 241 KTCPTNYMNNHICRCLAODEFMFSSDAGDSTDFGHDICGPNKELDEETCCQVCRAGLR 300
QY 301 PASGCPHRELDRNSQCVCCKNKLFPSCGANEPEDENTCCQCKRTCPRNOPLNPGKAC 360
DB 301 PASGCPHRELDRNSQCVCCKNKLFPSCGANEPEDENTCCQCKRTCPRNOPLNPGKAC 360
QY 361 ECTESPQCKLLGKKFHHOTCSYRRPCTNRQKACEPGFSYSEEVCRVPSYWRPQMS 419
DB 361 ECTESPQCKLLGKKFHHOTCSYRRPCTNRQKACEPGFSYSEEVCRVPSYWRPQMS 419

```

RESULT 5

```

US-09-042-105-18
Sequence 18, Application US/09042105
Patent No. 6040157

```

GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,105
FILING DATE: HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550
FILING DATE: 8-MAR-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TO BE ASSIGNED
FILING DATE: 24-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: ERIC K. STERRE
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-042-105-18

Query Match 99.6%; Score 2326; DB 3; Length 419;
Best Local Similarity 99.5%; Pred. No. 1,3e-201;
Matches 417; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MHLGFFSVACSLAAALLPGPREAPAAAAAFESGLDSDAEPDAGEATAYASKDLEQL 60
DB 1 MHSLEFFSVACSLAAALLPGPREAPAAAAAFESGLDSDAEPDAGEATAYASKDLEQL 60
QY 61 RSVSSVDELMATVLPYPMYMKCOLRKGMQHNREOANINSTEETIKFAAAHYNTIELK 120
DB 61 RSVSSVDELMATVLPYPMYMKCOLRKGMQHNREOANINSTEETIKFAAAHYNTIELK 120
QY 121 STDNEMRKQCPREVCDVKGKEFGVATNTFFKPCVSVYRCGGCNSGLOCMNTSTSY 180
DB 121 STDNEMRKQCPREVCDVKGKEFGVATNTFFKPCVSVYRCGGCNSGLOCMNTSTSY 180
QY 181 LSKTLFEITVPLSGPKPYTISFANHTSCRMKSLDVYRQVHSIIRSLPATLPQOQAN 240
DB 181 LSKTLFEITVPLSGPKPYTISFANHTSCRMKSLDVYRQVHSIIRSLPATLPQOQAN 240
QY 241 KTCPTNYMNNHICRLAEDFMFSSDAGDSDTGFHDICGPNKLEDETCOCVCACRAG 300
DB 241 KTCPTNYMNNHICRLAEDFMFSSDAGDSDTGFHDICGPNKLEDETCOCVCACRAG 300

DB 241 KTCPTNYMNNHICRLAEDFMFSSDAGDSDTGFHDICGPNKLEDETCOCVCACRAG 300
QY 301 PASCGPHKELDRNSCQCVCCKNKLFPSCGANCNPEDNTQCCKRCPNPNPLNPKCAC 360
DB 301 PASCGPHKELDRNSCQCVCCKNKLFPSCGANCNPEDNTQCCKRCPNPNPLNPKCAC 360
QY 361 ECTSPQKCLLGGKKRHHQTCSCYRPPCTNRKACPEGSYSEECVCPYSTWKRPQMS 419
DB 361 ECTSPQKCLLGGKKRHHQTCSCYRPPCTNRKACPEGSYSEECVCPYSTWKRPQMS 419

RESULT 6
PCT-US96-09001-2
Sequence 2, Application PC/TUS9609001
GENERAL INFORMATION:
APPLICANT: HU, ET AL.
TITLE OF INVENTION: Human Vascular Endothelial Growth Factor 2
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/09001
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/465,968
FILING DATE: 6 JUN 95
APPLICATION NUMBER: 08/207,550
FILING DATE: 8 MAR 1994
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-288
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
PCT-US96-09001-2

Query Match 99.3%; Score 2320; DB 5; Length 419;
Best Local Similarity 99.3%; Pred. No. 4.6e-201;
Matches 416; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MHLGFFSVACSLAAALLPGPREAPAAAAAFESGLDSDAEPDAGEATAYASKDLEQL 60
DB 1 MHSLEFFSVACSLAAALLPGPREAPAAAAAFESGLDSDAEPDAGEATAYASKDLEQL 60
QY 61 RSVSSVDELMATVLPYPMYMKCOLRKGMQHNREOANINSTEETIKFAAAHYNTIELK 120
DB 61 RSVSSVDELMATVLPYPMYMKCOLRKGMQHNREOANINSTEETIKFAAAHYNTIELK 120
QY 121 STDNEMRKQCPREVCDVKGKEFGVATNTFFKPCVSVYRCGGCNSGLOCMNTSTSY 180
DB 121 STDNEMRKQCPREVCDVKGKEFGVATNTFFKPCVSVYRCGGCNSGLOCMNTSTSY 180

LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-08-510-133A-33

Query Match 85.6%; Score 1999; DB 4; Length 350;
Best Local Similarity 100.0%; Pred. No. 2,9e-172;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 MVLVPEYKMKYKCOLRKGGWQHNRQANLNSRTEETIKFAAAHYNTILKSIDNEMRKT 129
DB 1 MVLVPEYKMKYKCOLRKGGWQHNRQANLNSRTEETIKFAAAHYNTILKSIDNEMRKT 60
QY 130 QCMPEVVCIDVKEGEGVANTFEKPPCVSVYRGCGGCSNSEGLOCNMTSTYSTLTFEIT 189
DB 61 QCMPEVVCIDVKEGEGVANTFEKPPCVSVYRGCGGCSNSEGLOCNMTSTYSTLTFEIT 120
QY 190 VPLSGPKPVITISFANHSCRCMSKLDVYRQVHSIIRSLPATLPQCOANKTCPTNYMM 249
DB 121 VPLSGPKPVITISFANHSCRCMSKLDVYRQVHSIIRSLPATLPQCOANKTCPTNYMM 180
QY 250 NNHICRCLAQEDPFMFSSDAGDSTDFHDIQCPNKELBDETCOCVCRAGLRPASCGRPHE 309
DB 181 NNHICRCLAQEDPFMFSSDAGDSTDFHDIQCPNKELBDETCOCVCRAGLRPASCGRPHE 240
QY 310 LDRNSQCVCKNKLFPSCGANREPDENTCQCVCKRTCPRNQPLNPGKACACETESPQK 369
DB 241 LDRNSQCVCKNKLFPSCGANREPDENTCQCVCKRTCPRNQPLNPGKACACETESPQK 300
QY 370 LKGGKFHHQTCSCYRRPCTNRQACEPFSYSEEVCRCPVSYMKRPQMS 419
DB 301 LKGGKFHHQTCSCYRRPCTNRQACEPFSYSEEVCRCPVSYMKRPQMS 350

RESULT 9
US-08-585-895-33

; Sequence 33, Application US/08585895

; Patent No. 6245530

; GENERAL INFORMATION:

; APPLICANT: Aitalo, Karl

; APPLICANT: Joukov, Vladimir

; TITLE OF INVENTION: Receptor Ligand

; NUMBER OF SEQUENCES: 35

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum

; STREET: 6300 Sears Tower, 233 South Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: United States of America

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/585,895

; FILING DATE:

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Gass, David A.

; REGISTRATION NUMBER: 38,153

; REFERENCE/DOCKET NUMBER: 28113/33072

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312/474-6300

; TELEFAX: 312/474-0448

; INFORMATION FOR SEQ ID NO: 33:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 350 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-585-895-33

Query Match 85.6%; Score 1999; DB 4; Length 350;
Best Local Similarity 100.0%; Pred. No. 2,9e-172;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 MVLVPEYKMKYKCOLRKGGWQHNRQANLNSRTEETIKFAAAHYNTILKSIDNEMRKT 129
DB 1 MVLVPEYKMKYKCOLRKGGWQHNRQANLNSRTEETIKFAAAHYNTILKSIDNEMRKT 60
QY 130 QCMPEVVCIDVKEGEGVANTFEKPPCVSVYRGCGGCSNSEGLOCNMTSTYSTLTFEIT 189
DB 61 QCMPEVVCIDVKEGEGVANTFEKPPCVSVYRGCGGCSNSEGLOCNMTSTYSTLTFEIT 120
QY 190 VPLSGPKPVITISFANHSCRCMSKLDVYRQVHSIIRSLPATLPQCOANKTCPTNYMM 249
DB 121 VPLSGPKPVITISFANHSCRCMSKLDVYRQVHSIIRSLPATLPQCOANKTCPTNYMM 180
QY 250 NNHICRCLAQEDPFMFSSDAGDSTDFHDIQCPNKELBDETCOCVCRAGLRPASCGRPHE 309
DB 181 NNHICRCLAQEDPFMFSSDAGDSTDFHDIQCPNKELBDETCOCVCRAGLRPASCGRPHE 240
QY 310 LDRNSQCVCKNKLFPSCGANREPDENTCQCVCKRTCPRNQPLNPGKACACETESPQK 369
DB 241 LDRNSQCVCKNKLFPSCGANREPDENTCQCVCKRTCPRNQPLNPGKACACETESPQK 300
QY 370 LKGGKFHHQTCSCYRRPCTNRQACEPFSYSEEVCRCPVSYMKRPQMS 419
DB 301 LKGGKFHHQTCSCYRRPCTNRQACEPFSYSEEVCRCPVSYMKRPQMS 350

RESULT 10

US-08-999-811-4

; Sequence 4, Application US/08999811

; Patent No. 5932540

; GENERAL INFORMATION:

; APPLICANT: HU, JING-SHAN

; APPLICANT: ROSEN, CRAIG A.

; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX

; STREET: 1100 NEW YORK AVENUE

; CITY: WASHINGTON

; STATE: DC

; COUNTRY: USA

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/999,811

; FILING DATE: HERewith

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/207,550

; FILING DATE: 8-MAR-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/465,968

; FILING DATE: 06-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: MARKOWICZ, KAREN R.

; REGISTRATION NUMBER: 36,351

; REFERENCE/DOCKET NUMBER: 1488.1000004

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)371-2600


```

Query Match      85.4%; Score 1995; DB 3; Length 350;
Best Local Similarity 99.7%; Pred. No. 6.7e-172;
Matches 349; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 70 MVLVPEYWMYKQQLRKGMQHNREOANLNSTEEIKFAAAHYNTIELKSIDNEKRKT 129
DB 1 MVLVPEYWMYKQQLRKGMQHNREOANLNSTEEIKFAAAHYNTIELKSIDNEKRKT 60
QY 130 QCMPEKVIDGKEFGVATNTFFKPCVSVYRGGCCNSGLQCMNTSTYLSKTLFEIT 189
DB 61 QCMPEKVIDGKEFGVATNTFFKPCVSVYRGGCCNSGLQCMNTSTYLSKTLFEIT 120
QY 190 VPLSGPKPVYTSIFANHTSCRCMSKIDVYRQVHSIIRSLPATLPOCAANKTCPTNYW 249
DB 121 VPLSGPKPVYTSIFANHTSCRCMSKIDVYRQVHSIIRSLPATLPOCAANKTCPTNYW 180
QY 250 NNHICRCLAOEDFMFSSDAGDDSTDPHDCGPNKELDEETCCQVCRAGLRPASCGRHKE 309
DB 181 NNHICRCLAOEDFMFSSDAGDDSTDPHDCGPNKELDEETCCQVCRAGLRPASCGRHKE 240
QY 310 LDRNSQCVCYKNTLFPSCGAGNREFDENTCCQVCKRTCPNQPPLNPGKACCECTESPQK 369
DB 241 LDRNSQCVCYKNTLFPSCGAGNREFDENTCCQVCKRTCPNQPPLNPGKACCECTESPQK 300
QY 370 LKGGKFFHQTCSCYRRPCTNRQKACEPGFSYSEEVCRCPVSYWKRQMS 419
DB 301 LKGGKFFHQTCSCYRRPCTNRQKACEPGFSYSEEVCRCPVSYWKRQMS 350

RESULT 13
US-08-795-430-13
; Sequence 13, Application US/08795430
; Patent No. 6130071
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
; TITLE OF INVENTION: Protein and Gene, Mutants Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,430
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP196/00427
; FILING DATE: 01-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/671,573
; FILING DATE: 28-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/601,132
; FILING DATE: 14-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/585,895
; FILING DATE: 12-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/510,133
; FILING DATE: 01-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,011
; FILING DATE: 14-NOV-1994

```

```

ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 28967/33691
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 418 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-795-430-13

Query Match      77.2%; Score 1804.5; DB 4; Length 418;
Best Local Similarity 75.7%; Pred. No. 1.2e-154;
Matches 318; Conservative 39; Mismatches 60; Indels 3; Gaps 3;

QY 1 MHLGFFSVACSLAALLDGPREAPAAAAPFSSGLDSDAEPDAGEATAYASKLEBOL 60
DB 1 MHLGFFSVACSLAALLDGPREAPAAAAPFSSGLDSDAEPDAGEATAYASKLEBOL 59
QY 61 RSVSYDELMTVLYPEYWMYKQQLRKGMQHNREOANLNSTEEIKFAAAHYNTIELK 120
DB 60 RSVSYDELMTVLYPEYWMYKQQLRKGMQHNREHSSDTRSDSLKFAAAHYNTIELK 119
QY 121 SIDNEKRKTQCMPEKVIDGKEFGVATNTFFKPCVSVYRGGCCNSGLQCMNTSTY 180
DB 120 SIDNEKRKTQCMPEKVIDGKEFGVATNTFFKPCVSVYRGGCCNSGLQCMNTSTY 179
QY 181 LSKTLFEITVPLSGPKPVYTSIFANHTSCRCMSKIDVYRQVHSIIRSLPATLPOCAAN 240
DB 180 LSKTLFEITVPLSGPKPVYTSIFANHTSCRCMSKIDVYRQVHSIIRSLPATLPOCAAN 239
QY 241 KICPTVYMNNNHICRCLAOEDFMFSSDAGD-DSTDPHDCGPNKELDEETCCQVCRAGL 299
DB 240 KICPTVYMNNNHICRCLAOEDFMFSSDAGD-DSTDPHDCGPNKELDEETCCQVCRAGL 298
QY 300 RPASGPKHKLDRNSQCVCYKNTLFPSCGAGNREFDENTCCQVCKRTCPNQPPLNPGKCA 359
DB 299 RPASGPKHKLDRNSQCVCYKNTLFPSCGAGNREFDENTCCQVCKRTCPNQPPLNPGKCA 358
QY 360 CECTESPQKCLKGGKFFHQTCSCYRRPCTNRQKACEPGFSYSEEVCRCPVSYWKRQMS 419
DB 359 CECTESPQKCLKGGKFFHQTCSCYRRPCTNRQKACEPGFSYSEEVCRCPVSYWKRQMS 418

RESULT 14
US-08-915-795-3
; Sequence 3, Application US/08915795
; Patent No. 6235713
; GENERAL INFORMATION:
; APPLICANT: Marc G. ACHEN
; APPLICANT: Andrew F. WILKS
; APPLICANT: Steven A. SPACKER
; APPLICANT: Kari ALITALO
; TITLE OF INVENTION: GROWTH FACTOR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
; STREET: 1200 G Street, NW, Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: United States of America
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25

```

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/915,795
 FILING DATE:
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: EVANS, Joseph D.
 REGISTRATION NUMBER: 26,269
 REFERENCE/DOCKET NUMBER: 1064/429830
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-8800
 TELEFAX: (202) 628-8844
 TELEX: N/A
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 325 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEetical: NO
 ORIGINAL SOURCE:
 TISSUE TYPE: Human Breast

Query Match	30.2%	Score 704.5;	DB 4;	Length 325;
Best Local Similarity	38.9%	Pred No. 9.8e-56;		
Matches 140; Conservative	61;	Mismatches 88;	Indels 71;	Gaps 11;

QY	57	EEOLRSYSDVELMTVTVYPERMKYKCOLKRGQWONHREOANLSR--TEETIKFAAHY	114
Db	12	EQOIRAAASLEELLRIHSEDMKLMRCRLR-----KSFSDMSRSASHRSIRFANLEY	65
QY	115	MTIELKSLDNRKTCQCMRPVCLDYGKEGVAVNTFFKPCVSVYKCGGCCSEGIQCM	174
Db	66	DETLKTVDEMOORTQSPRTCEVAISELGSINTFFPKPCVAVFQCGGCCSEELICM	125
QY	175	NTSTYLSKLTLEITVPLSQGPKRVITISFANTHSCRMKSLDYVRQVHSTIRRSPLTP	234
Db	126	NTSTYISKQLEFELSVPLTSVELVPAKVANHHCCKLPLAP--RHPIYSIIRSI--QIP	181
QY	235	Q---QQAANKTCSPNNYMMNNHHCGLAODEPMFSMDAGDSDTGDFHDTGCPNKLEDETC	291
Db	182	EEDRSKSHKTCSPIDMLMDSMKKCCVLOE-----ENPLAGTED---HSHLOE--	225
QY	292	QCVCRAGLRAPACGPKHELDNRSCOCVSKNKLFPSCGAGAREDEMTQOCVCRTSPRNO	351
Db	226	-----PALGSPH-----MMDEDEGCEVCTTPCKDL	252
QY	352	PLNPGKACAC--ECTSSPOKCLLKGKRYHNQTC-----YRRPCTNRQAKCEGFSTSEE	404
Db	253	IOHNRNSCECFKSKSLTSCQKNLTFHDDTSCEDRRPCTNRQASCKTCAAHNRPKPE	312

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/915,795
 FILING DATE:
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: EVANS, Joseph D.
 REGISTRATION NUMBER: 26,269
 REFERENCE/DOCKET NUMBER: 1064/42983
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-8800
 TELEFAX: (202) 628-8844
 TELETYPE: N/A
 INFORMATION FOR SEO ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 354 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 TISSUE TYPE: Human Lung

Query Match	30.2%	Score 704.5	DB 4	Length 354
Best Local Similarity	38.9%	Pred. No. 1.1e-55		
Matches 140; Conservative	61;	Mismatches 88;	Indels 71;	Gaps 11;

Qy	57	EOELRSVDELMTVLYPEYKMYKQOLRRGQOHNRQANLSNR--TEETIEFAAHY	114
Db	41	EQOLRASSLEELRIHSEBMKIMCRRLR-----KSPISMDRSASHRSTFEATFY	94
Qy	115	NTIELKSIDNEMKRTQCMPREVCIIDVCKEKGAVANTTFPKPCVSVYRGGCCNSGLQM	174
Db	95	DIELFKYIDEMOHTQCSPRETCVEVASELSKSNTEFFKPPCVAVYFPGGCGNSESILCM	154
Qy	175	NTSTSYLKTLEETVTPLSQGRPVITISFANTHSCRCMSKIDVYRQVHSIIRSLPATLP	234
Db	155	NTSTSYISKQLEISSLVPLTSVELYPAVKVANHCKGLPLAP--RHXYSIIRRI--QIP	210
Qy	235	Q---QQAANKTCPTNYMMNNHTICCLAOEDBFMSSDAGDSDTSCFHICGPNKELDETC	291
Db	211	EEDRCSHSKTKCPIIDMLTMSNKKCKCVLOE-----ENPLAOTED---HSHLOE--	254
Qy	292	QCVCYRAGLRPAQCPNHELDNRNSCQCVCNKLFPSCGAGNREPEMTCQCCKYKTRIRNQ	351
Db	255	-----PALGCPH-----MMFEDRCBECVCKPCPKDL	281
Qy	352	PLNKGKACG-ECTSSPOKCLLKGGKHHNHTSC-----YRRPCTNQAACEPFSISEE	404
Db	282	IQHKNKSCFECKSKSLTECCQKHNLEFPDTCSCDRCPIFHTRPASGSKTCAKHKRPKE	341

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:43:54 ; Search time 78.16 Seconds
(without alignments)
408.356 Million cell updates/sec

Title: US-09-427-657-2

Perfect score: 2336

Sequence: 1 MHLLGFSSVACSLIAALLP.....SYSEVVCRCVSYWKRPQMS 419

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2336	100.0	419	2	6S9207
2	257	11.0	1700	2	S08167
3	238.5	10.2	232	2	M41551
4	218.5	9.4	190	2	S52130
5	207.5	8.9	190	2	B44881
6	207.5	8.9	214	2	A44881
7	206.5	8.8	190	2	A35987
8	205.5	8.8	190	2	B40080
9	182.5	7.8	473	2	A56175
10	180.5	7.7	160	2	U00542
11	175	7.5	188	2	JC4680
12	173.5	7.4	146	2	S57956
13	172.5	7.4	120	2	A33787
14	172.5	7.4	1810	1	A32230
15	160	6.8	148	2	D95530
16	159.5	6.8	133	2	B49530
17	158	6.8	2703	1	A24420
18	157.5	6.7	1746	1	S19694
19	157.5	6.7	2471	2	A49128
20	157	6.7	3635	2	T10053
21	155.5	6.7	2918	2	A4105
22	155	6.6	1220	2	A56136
23	154.5	6.6	2907	2	A57278
24	154	6.6	1372	2	T25933
25	154	6.6	2437	2	T30201
26	153.5	6.6	2352	2	A5844
27	152.5	6.5	2524	2	A5844
28	151.5	6.5	782	2	A61625
29	151	6.5	149	2	A41236

30	150.5	6.4	1203	2	A49175
31	149.5	6.4	1187	2	T18355
32	149	6.4	2201	2	A32160
33	149	6.4	2871	2	A55624
34	149	6.4	3002	2	A47221
35	148.5	6.4	207	2	JC4679
36	148.5	6.4	1871	2	A5567
37	147.5	6.3	1531	2	T42218
38	147.5	6.3	1964	2	T09059
39	147.5	6.3	2321	2	S78549
40	146	6.2	2531	2	T31070
41	145.5	6.2	565	2	T16408
42	145	6.2	1574	2	T13954
43	144	6.2	158	2	A56125
44	144	6.2	2813	1	VWNU
45	144	6.2	4006	2	T09070

ALIGNMENTS

RESULT 1
S69207
vascular endothelial growth factor C precursor - human
N:Alternate names: FLT4 ligand DHM
C:Species: Homo sapiens (man)
C>Date: 27-Apr-1996 #sequence_revision 01-Nov-1996 #text_change 08-Oct-1999
C:Accession: S69207; S61795; S71443; S69208; G02659
R:Joukov, V.; Pajusola, K.; Kaipainen, A.; Chilov, D.; Laitinen, I.; Kukk, E.; Saksel
EMBO J. 15, 1751, 1996
A>Title: Corrigendum: A novel vascular endothelial growth factor, VEGF-C, is a ligand
A:Reference number: S69207; MUID:96203094
A:Accession: S69207
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-419 <JOU>
A:Cross-references: EMBL:X94216; NID:q1177488; PIDN:CAA63907.1; PID:e221096; PID:g118
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1995
A>Note: only a part of the translation is shown
A>Note: this is a revision to the sequence from reference S61795
R:Joukov, V.; Pajusola, K.; Kaipainen, A.; Chilov, D.; Laitinen, I.; Kukk, E.; Saksel
EMBO J. 15, 290-298, 1996
A>Title: A novel vascular endothelial growth factor, VEGF-C, is a ligand for the FLT4
A:Reference number: S61795; MUID:96178224
A:Accession: S61795
A>Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 70-419 <JOU1>
A:Note: this sequence has been revised in reference S69207
A:Accession: S71443
A:Molecule type: protein
A:Residues: X, 104-120 <JOU2>
R:Lee, J.; Gray, A.; Yuan, J.; Luoh, S.M.; Avraham, H.; Wood, W.I.
submitted to the EMBL Data Library, December 1995
A:Description: Vascular endothelial growth factor related protein (VRP): A ligand and
A:Reference number: S69208
A:Accession: S69208
A:Molecule type: mRNA
A:Residues: 1-419 <LEE>
A:Cross-references: EMBL:U43142; NID:g1150988; PIDN:AAA5214.1; PID:g1150989
R:Morris, J.C.
submitted to the EMBL Data Library, May 1996
A:Reference number: H01557
A:Accession: G02659
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-419 <MOR>
A:Cross-references: EMBL:U58111; NID:g1373426; PIDN:ABD02909.1; PID:g1373427
C:Genetics:
A:Gene: GDB:VEGFC; VRP
A:Cross-references: GDB:3890883; OMIM:601528
F:1-12/Domain: signal sequence #status predicted <SIG>
F:13-102/Domain: propeptide #status predicted <PRO>

Notch B protein -
hypothetical prote
tenascin-C - human
fibritin-1 precu
fibritin-1 precu
vascular endotheli
fibritin-1 - bovi
silt-1 protein hom
notch4 - mouse
notch3 protein - h
hypothetical prote
MEGF6 protein - ra
placental growth fac
von Willebrand fac
probable tenascin

F:103-419/Product: vascular endothelial growth factor C #status experimental <MAY>

Query Match 100.0%; Score 2336; DB 2; Length 419;

Best Local Similarity 100.0%; Pred. No. 1,3e-164;

Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 MHLGFFVASCILAAALLPGPREAPAAAAAFESGLDLSADAPDAGEAFAVASKDEBL 60
Db 1 MHLGFFVASCILAAALLPGPREAPAAAAAFESGLDLSADAPDAGEAFAVASKDEBL 60
Oy 61 RSVSYVDLMTVLYPEYKMKYKQQLRKGWQHNRQOANLSRTETIKFAAHYNTIELK 120
Db 61 RSVSYVDLMTVLYPEYKMKYKQQLRKGWQHNRQOANLSRTETIKFAAHYNTIELK 120
Oy 121 STDNERKRTQCPREYCIDVGEFVATNTFFKPCVSVYRGGCCNSGLQCMNTSTY 180
Db 121 STDNERKRTQCPREYCIDVGEFVATNTFFKPCVSVYRGGCCNSGLQCMNTSTY 180
Oy 181 LSKTLEITVPSOGPKPVTISFANHTSCRCMSKLDVYRQVHSIIRSLPATLPQOQAN 240
Db 181 LSKTLEITVPSOGPKPVTISFANHTSCRCMSKLDVYRQVHSIIRSLPATLPQOQAN 240
Oy 241 KTCPTNYMNNHICRCLAQEDFMFSSDAGDDSTDFHICGPNKELDEFTCCQVCRAGLR 300
Db 241 KTCPTNYMNNHICRCLAQEDFMFSSDAGDDSTDFHICGPNKELDEFTCCQVCRAGLR 300
Oy 301 PASCGHKLDRNSCQCCVCKNKLFPSCGANCNEFDNTCCQCKRCPCPNQPLNPKKAC 360
Db 301 PASCGHKLDRNSCQCCVCKNKLFPSCGANCNEFDNTCCQCKRCPCPNQPLNPKKAC 360
Oy 361 ECTESPQKLLGKGFHHTQSCYRRPCTNRKACPEGSYSSEVCRCPYSWKRPQMS 419
Db 361 ECTESPQKLLGKGFHHTQSCYRRPCTNRKACPEGSYSSEVCRCPYSWKRPQMS 419

RESULT 2
S08167
Balblani ring 3 protein - midge (Chironomus tentans)
C:Species: Chironomus tentans
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Jul-2000
C:Accession: S08167
R:Paulsson, G.; Lendahl, U.; Galli, J.; Ericsson, C.; Wieslander, L.
J. Mol. Biol. 211, 331-349, 1990
A:Title: The balblani ring 3 gene in Chironomus tentans has a diverged repetitive struct
A:Reference number: S08167; MUID:90172404
A:Accession: S08167
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-1700 <PAUV>
A:Cross-references: GB:X52263; NID:g7057; PIDN:CAA36506.1; PID:g7058
C:Genetics:
A:Gene: BR3
A:Map position: 4
C:Superfamily: unassigned Balblani ring proteins

Query Match 11.0%; Score 257; DB 2; Length 1700;
Best Local Similarity 22.8%; Pred. No. 5,1e-11;
Matches 89; Conservative 47; Mismatches 145; Indels 110; Gaps 18;

Oy 79 KMYKCOLRK---GWMQHNRQOANLSRTETIKFAAHYNTIELKSIDNERKRTQCMR 134
Db 727 KTKCKCEKEMPTGCGENKMKC-----DETCDCVCQKNTCIAPKV---WDAKTC-- 775
Oy 135 EVCIDVGEFVATNTFFKPCVSVYRGG-----CCNSEGLQCMNTSTYSLKTLF 186
Db 776 --CI-----CVNPPCKSPQVLDKTCGCGQANVSKCAPOKFT-ENIC 815
Oy 187 EITVPLSGPKPVTISFANHTSCRCMSKLDV-----YRQVHSIIRSLPATLPQOQA 238
Db 816 DQACPRKCKCKAPLYWSDEFCDCVCPNSASMTCLSPKEMKVTCTCDCCNPK--PDC-- 871

```

```

Oy 239 ANKTCPTNYMNNHICRC---LAQEDFMFSSDAGDDSTDFHICG-PNKELD----- 287
Db 872 ----CGTQKMMWDDKCKCGCPNAQTDC-----AGKKRFNDFTCSCGCPBGKIDCTGNMKM 922
Oy 288 -EETCCVCRAGLRPAPSKPHKELDRNSCQVCKNKLFPSCGANCNEFDNTCCQVCR- 345
Db 923 SAETCTCGC--GDVNRNCGNLKFNFDNLQCCCKKQKQKOEANCKSPRTWNYDCKKVCCKNA 980
Oy 346 -----TCPRNQPLNPGKACGECIESPQKLLKGRKFFH 378
Db 981 DDSDCVKQIWLDDQCKCGCPASQMTCPANKRFTESCSCEKSPMPSPPIQCKMWE 1040
Oy 379 QTCSCYRRPCTNRKACPEGSYSSEVCRCV 409
Db 1041 DKCVV---ECAN-VKTCBGPQRWCDNQCKCI 1067

RESULT 3
A41551
vascular endothelial growth factor 206 precursor - human
M:Alternate names: vascular permeability factor
N:Comments: vascular endothelial growth factor 121 (VEGF 121); VEGF 165; VEGF 189; VE
C:Species: Homo sapiens (man)
C:Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 05-Nov-1999
C:Accession: A41551; C41551; B41551; A40454; B40454; C40454; A40079; A40080; JQ1463;
R:Honck, K.A.; Ferrara, N.; Winer, J.; Cachianes, G.; Li, B.; Leung, D.W.
Mol. Endocrinol. 5, 1806-1814, 1991
A:Title: The vascular endothelial growth factor family: identification of a fourth mo
A:Reference number: A41551; MUID:92168017
A:Accession: A41551
A:Molecule type: mRNA
A:Residues: 1-232 <HOU1>
A:Cross-references: GB:S85192; NID:g246155; PID:g246156
A:Accession: C41551
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-140, 'N', 183-232 <HOU2>
A:Accession: B41551
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-141, 227-232 <HOU>
R:Tischer, E.; Mitchell, R.; Hattman, T.; Silva, M.; Gospodarowicz, D.; Fiddes, J.C.;
J. Biol. Chem. 266, 11947-11954, 1991
A:Title: The human gene for vascular endothelial growth factor. Multiple protein form
A:Reference number: A40454; MUID:91268072
A:Accession: A40454
A:Molecule type: DNA
A:Residues: 1-165, 183-232 <TI1>
A:Cross-references: GB:M63972; GB:M63973; GB:M63974; GB:M63975; GB:M63976;
A:Accession: B40454
A:Molecule type: DNA
A:Residues: 1-140, 'N', 183-232 <TI2>
A:Cross-references: GB:M63971; GB:M63972; GB:M63973; GB:M63974; GB:M63975; GB:M63977;
A:Accession: C40454
A:Molecule type: DNA
A:Residues: 1-141, 227-232 <TI3>
A:Cross-references: GB:M63971; GB:M63972; GB:M63973; GB:M63974; GB:M63975; GB:M63978
R:Keck, P.J.; Hauser, S.D.; Krivi, G.; Sanzo, K.; Warren, T.; Feder, J.; Connolly, D.
Science 246, 1309-1312, 1989
A:Title: Vascular permeability factor, an endothelial cell mitogen related to PDGF.
A:Reference number: A40079; MUID:90069609
A:Accession: A40079
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-165, 183-232 <KEC>
A:Cross-references: GB:M27261; NID:g340300; PIDN:AAA36807.1; PID:g340301
R:Leung, D.W.; Cachianes, G.; Kuang, W.J.; Goeddel, D.V.; Ferrara, N.
Science 246, 1306-1309, 1989
A:Title: Vascular endothelial growth factor is a secreted angiogenic mitogen.
A:Reference number: A40080; MUID:90069608
A:Accession: A40080
A:Status: not compared with conceptual translation
A:Molecule type: mRNA

```

A:Residues: 1-140,'N',183-232 <LEU>
A:Cross-references: GB:M32977; NID:g181970; PIDN:AAA35789.1; PID:g181971
R:Weinandel, K.; Matme, D.; Welch, H.A.
Biochem. Biophys. Res. Commun. 183, 1167-1174, 1992
A:Title: AIDS-associated Kaposi's sarcoma cells in culture express vascular endothelial
A:Reference number: J01463; MUID:92231879
A:Accession: J01463
A:Molecule type: mRNA
A:Residues: 1-140,'N',183-232 <WEI>
A:Cross-references: EMBL:X62568; NID:937658; PIDN:CAA44447.1; PID:g37659
A:Experimental source: AIDS-Kaposi's sarcoma cell
A:Accession: J01464
A:Molecule type: mRNA
A:Residues: 1-140,'N',227-232 <WE2>
A:Experimental source: AIDS-Kaposi's sarcoma cell
R:Connolly, D.T.; Olander, J.V.; Heuvelman, D.; Nelson, R.; Monsell, R.; Siegel, N.; Hay
J. Biol. Chem. 264, 20017-20024, 1989
A:Title: Human vascular permeability factor. Isolation from U937 cells.
A:Reference number: A34492; MUID:90062112
A:Accession: A34492
A:Molecule type: protein
A:Residues: 27-36;43-49,'R',72-76,'Q',78-81;59-71 <CON>
A:Comment: The most common of several alternatively spliced forms is VEGF 165.
C:Genetics:
A:Gene: GDB:VEGF
A:Cross-references: GDB:132244; OMIM:192240
A:Map position: 6p21-6p12
C:Function:
A:Description: promotes fluid and protein leakage from blood vessels
C:Keywords: alternative splicing; angiogenesis; dimer; disulfide bond; extracellular prote
F1-1-332/Product: vascular endothelial growth factor 206 precursor #status predicted <Y0
F1-1-165,183-233/Product: vascular endothelial growth factor 189 precursor #status predi
F1-1-141,227-233/Product: vascular endothelial growth factor 121 precursor #status predi
F1-1-26/Domains: signal sequence #status predicted <SIG>
F1-1-6/Binding site: carbohydrate (Asn) (covalent) #status predicted

[illegible]

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-190 <SHA>
A:Cross-references: GB:X81380; NID:9587559; PIDD:CAA57143.1; PID:9587560

	Query Match	9.4%	Score 218.5;	DB 2;	Length 190;	
	Best Local Similarity	23.3%	Pred. No. 4e-09;			
	Matches	62;	Conservative	25;	Indels 107;	Gaps 8.
Oy	18KGGQHNREDQANLSRTEITIKFAAHHYNTELKSIDNEMRKTCQMREVCIDVGRKEF	144				
Dd	18 LHAHAKSOAPRAEEDQKHVEYVKM-----DYVORSTCRPTFLVDIYDFCY	64				
Oy	145 GVAINTPEKKPCVSUYRCGCGSCNEBGLQCMNTSTYSKLTEITIVPL-----SOGPKFY	199				
Dd	65 PDEIETIRKPSCVPLMRGCGGCCNDGBLEVPTEE-----FNITMQIMRIKPHQGNGH	117				
Oy	200 TISFANHSTSCRMSKLDVYRVYHSIIIRSLPATLRPOQAANKTCTPTNYMMNNHGICSLAQ	259				
Dd	118 EMSFLQHNKCECRPKKRDRARQ-----	138				
Oy	260 EDFMSSDAGDSDTGCFHDICGPKNKEIDEETQCVCVRAGLPRASGCPKEL---DRNSCQ	316				
Dd	139 -----ENPGCGPGE-----RRKHLFVODPOTCK	161				
Oy	317 CYCKRKLFPPSGCANR--EPDENTQQC	341				
Dd	162 CSCSKNT--DSRCKARQLLETNETRCKC	185				

B44881
vascular endothelial growth factor-1 precursor - mouse

C:Date: 03-Feb1994 #sequence_revision 03-Feb-1994 #text_change 05-Nov-1999
C:Accession: B44881, A43351, A61029
R:Beier, G.; Albrecht, U.; Sterrer, S.; Risau, W.
Development 114, 521-532, 1992
A:Title: Expression of vascular endothelial growth factor during embryonic angiogenesis
A:Reference number: A44881; MUID:92274860
A:Accession: B44881
A:Molecule type: mRNA
A:Residues: 1-190 <BRE>
A:Cross-references: GB:S38083; NID:g249858; PIDN:AAB22253.1; PID:g249859
A:Experimental source: embryo
R:Claffey, K.P.; Wilkison, W.O.; Spiegelman, B.M.
J. Biol. Chem. 267, 16317-16322, 1992
A:Title: Vascular endothelial growth factor. Regulation by cell differentiation and a
A:Reference number: A43351; MUID:92355593
A:Accession: A43351
A:Molecule type: mRNA
A:Residues: 1-116, ER_119-190 <CLA>
A:Cross-references: GB:g95200; NID:g202350; PIDN:AAA0547.1; PID:g202351
R:Rosenthal, R.A.; Megyesi, J.F.; Henzel, W.J.; Ferrara, N.; Folkman, J.
Growth Factors 4, 53-59, 1990
A:Title: Conditioned medium from mouse sarcoma 180 cells contains vascular endothelial
A:Reference number: A61029; MUID:91197543
A:Accession: A61029
A:Molecule type: protein
A:Residues: 27-38 <ROS>
A:Keywords: alternative splicing; angiogenesis; dimer; disulfide bond; glycoprotein;

Query Match	8.9%	Score 207.5	DB. 2	Length 190
Best Local Similarity	22.7%	Pred. No. 2:6e-06		
Matches 59	Conservative 28	Mismatches 76	Indels 95	Gaps 8
QY	85	L R K G W G H N R E A N I N S T E E T I K F A A H A Y N T E I L S I D N E M R K T Q C M P R E V C I D V G K E F	144	
db	18	L H A K W S O A A P T T E O K S S H E V I K F M - - - - - D V O R S C Y P I F L V D I F O E Y	64	

[illegible]

Db 101 GTFECVCRNNGFGRKLECKNWCSPNPKNNG-KC-----SPLGKTGYKCTCS----- 146

QY 200 TISFANHSCRCMSKLDVYROYHSIIRSLPALTLPO--COAANKTCP---TNYMNNNHICR 255

Db 147 -----GGYGRPC-----EVH-----ACKRNPKNNGRCRPPDOKTGYK-----CR 181

QY 256 CL-----AQEDPFP-----SSDAGDSDTDGFHD-----ICGPKKEIDE 288

Db 182 CVDGYSGPFTCKENACKPNPNCSSNGGTCSADKFGDYSCECRPGYFGDECEERYVCAPIPKNG 241

QY 289 ETC-----QCVRAGLGRPASC-----GPHKEIDR-----NSCOCVCKNKLFPSCQ 328

Db 242 GTCSSDGSGGYRCRCCKGGYSGPFTCKAVNYCKLPTPCKNNSGRGVNKGSSYNCICKGGYSGFTC 301

QY 329 GAN-----REPENT-----COCV-----CK-----RTCPRRQPLNPGKC----- 358

Db 302 GENNVCKPNPCCNKGRCYFDNSDDGFKCKCVGGYKGPCTEDKPNPCNTCKPCKNGGCKWTNG 361

QY 359 ---ACECT-----ESPQCKLKGKRF-----HHQTC 381

Db 362 KLYTCKAAYGMGRGHCTDKAYKPNPCVYSPKPCNKGKIMNGKAYRCCKAYGVGGRHCTK 421

QY 382 SCYR-----RECTNRQKACEPGEFSYSEVCRCPYSY 412

Db 422 KSYKKNPCASRPCCKNRGCKTCKGNGY---VCKCARGY 455

RESULT 10

J00542

185K secretory protein - midge (Chironomus tentans) (fragment)

N:Alternate names: balbiani ring 3 protein

C:Species: Chironomus tentans

C:date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 17-Mar-2000

C:Accession: J00542

R:Dignam, S.S.; Case, S.T.

Gene 88, 133-140, 1990

A:Title: Balbiani ring 3 in Chironomus tentans encodes a 185-kDa secretory protein wh

A:Reference number: J00542; MUID:90269600

A:Accession: J00542

A:Molecule type: mRNA

A:Residues: 1-160 <DTG>

A:Cross-references: GB:M24160

A:Experimental source: salivary gland

C:Superfamily: unassigned Balbiani ring proteins

Query Match 7.7%; Score 180.5; DB 2; Length 160;

Best Local Similarity 26.4%; Pred. No. 2.1e-06;

Matches 48; Conservative 24; Mismatches 77; Indels 33; Gaps 9;

QY 243 CPTNMNNNHICRCLAQEDPFPSSDAGDSDTDGFHDICGPKKEIDEETCCVCRAGLRPA 302

Db 3 CKSPRQWTDSKCLC-----ECSTTPATEGKQWGWG-----EACQCICTPGG--DK 45

QY 303 SCGPHKEIDRNSCOCVCKNKLFPSSOCANREFEDNTCCQCVCKRT-----CPNQPPLNP 355

Db 46 NCGNNKKFPDKSCCECKKNN--PST--SPQYWDADDECECKCPKQKQKRGSGCDDGQKWN 101

QY 356 GKACACECTESPOKCLLKGKRFHHQTCSCYRRPCTNRQKACEPGEFSYSEVCR--VPYSYK 414

Db 102 RVCSCGCVPRPDC-TNGQIYINVTAC--GCCIDRPSCKQOIYMWKTCDCBCPNGMK 157

QY 415 RP 416

Db 158 RP 159

RESULT 11

J04680

vascular endothelial growth factor-related factor 167 precursor - mouse

N:Alternate names: VRF 167 protein

C:Species: Mus musculus (house mouse)

EMBO J. 7, 2977-2982, 1988
A:Title: Tenascin: cDNA cloning and induction by TGF-beta.
A:Reference number: S01292; MUID:89030589
A:Accession: S01292
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 27-181, 'R', 183-204, 'G', 206-221, 'A', 223-380, 'D', 382-386, 'H', 388-444, 'NN', 447-458
A:Cross-references: EMBL:X08030
A:Note: part of this sequence was confirmed by protein sequencing
C:Superfamily: tenascin; EGF homology; fibronogen beta/gamma homology; fibronectin type
C:Keywords: alternative splicing; calcium binding; cell adhesion; duplication; extracel
E:23-33/Domain: signal sequence #status predicted <SIG>
E:34-1810/Product: tenascin 230K #status predicted <MAT>
E:223-249/Domain: EGF homology <EGF1>
E:316-342/Domain: EGF homology <EGF>
E:592-673/Domain: fibronectin type III repeat homology <FN3A>
E:681-765/Domain: fibronectin type III repeat homology <FN3B>
E:773-857/Domain: fibronectin type III repeat homology <FN3C>
E:865-949/Domain: fibronectin type III repeat homology <FN3D>
E:957-1037/Domain: fibronectin type III repeat homology <FN3E>
E:1046-1128/Domain: fibronectin type III repeat homology <FN3F>
E:1137-1210/Domain: fibronectin type III repeat homology <FN3G>
E:1228-1310/Domain: fibronectin type III repeat homology <FN3H>
E:1318-1399/Domain: fibronectin type III repeat homology <FN3I>
E:1407-1487/Domain: fibronectin type III repeat homology <FN3J>
E:1495-1575/Domain: fibronectin type III repeat homology <FN3K>
E:1590-1798/Domain: fibronogen beta/gamma homology <FBG>
E:1734-1747/Domain: calcium binding #status predicted <CAB>

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:43:54 ; Search time 42.92 Seconds

(without alignments)
334.414 Million cell updates/sec

Title: US-09-427-657-2

Perfect score: 2336

Sequence: 1 MHLLGFESVACSLIAALLP.....SYSEEVCRVPSYMKRPQMS 419

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2336	100.0	419	1	VEGC_HUMAN
2	2048	87.7	415	1	VEGC_MOUSE
3	257	11.0	1700	1	BAR3_CHITE
4	226	9.7	215	1	VEGF_HUMAN
5	218.5	9.4	190	1	VEGF_PIG
6	207.5	8.9	214	1	VEGF_MOUSE
7	206.5	8.8	190	1	VEGF_RAT
8	205.5	8.8	190	1	VEGF_BOVIN
9	199.5	8.5	164	1	VEGF_CAVPO
10	182.5	7.8	473	1	FP2_MYGA
11	178.5	7.6	216	1	VEGF_CHICK
12	175.5	7.5	188	1	VEGB_HUMAN
13	175	7.5	188	1	VEGB_MOUSE
14	173.5	7.4	146	1	VEGF_SHEEP
15	163.5	7.0	2482	1	VEGF_PIG
16	162.5	7.0	133	1	VEGH_OREN2
17	160.5	6.9	1808	1	TENA_CHICK
18	160	6.8	148	1	VEGH_OREN7
19	158	6.8	2703	1	NOTC_DROME
20	157.5	6.7	1746	1	TENA_PIG
21	157	6.7	3635	1	LMAS_MOUSE
22	156.5	6.7	2813	1	VWF_CANFA
23	155.5	6.7	2911	1	FN2_HUMAN
24	154.5	6.6	2907	1	FN2_MOUSE
25	154	6.6	2437	1	NOTC_BRARE
26	152.5	6.5	2524	1	NOTC_XENLA
27	150	6.4	170	1	PLGF_HUMAN
28	149	6.4	2201	1	TENA_HUMAN
29	149	6.4	2871	1	FN1_HUMAN
30	149	6.4	2871	1	FN1_MOUSE
31	148.5	6.4	2871	1	FN1_BOVIN
32	147.5	6.3	1964	1	FN4_MOUSE
33	147	6.3	2444	1	NTC1_HUMAN

34	144	6.2	2813	1	VWF_HUMAN	P04275	homo sapien
35	143	6.1	3672	1	LMC2_CAEL	Q21313	caenorhabdi
36	142.5	6.1	1106	1	SIC_DROME	P40798	drosophila
37	142	6.1	158	1	PLGF_MOUSE	P49764	mus muscicu
38	141.5	6.1	1429	1	L112_CAEL	P14585	caenorhabdi
39	141	6.0	931	1	EMR1_MOUSE	Q61549	mus muscicu
40	140	6.0	2531	1	NTC1_MOUSE	Q01705	mus muscicu
41	139	6.0	2139	1	CRB_DROME	P10040	drosophila
42	138	5.9	1696	1	PCK5_BRACL	Q9h315	branchiosto
43	137	5.9	5179	1	MGC2_HUMAN	Q02817	homo sapien
44	136.5	5.8	1104	1	NFX1_HUMAN	Q12986	homo sapien
45	135.5	5.8	2531	1	NTC1_RAT	Q07008	rattus norv

ALIGNMENTS

```

RESULT 1
VEGC_HUMAN          STANDARD:      PRT:      419 AA.
ID      VEGC_HUMAN          AC      P49767:
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      01-OCT-2000 (Rel. 40, Last annotation update)
DE      VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR (VEGF-C) (VASCULAR
DE      ENDOTHELIAL GROWTH FACTOR RELATED PROTEIN) (VRF) (FLT4 LIGAND) (FLT4-
DE      L).
GN      VEGFC.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_Taxid=9606;
RN      [1]
RP      SEQUENCE FROM N.A., AND SEQUENCE OF 103-120.
RX      MEDLINE=96178224; PubMed=8617204;
RA      Joukov V., Pajusola K., Kaipainen A., Chillov D., Lahtinen I., Kukk E.,
RA      Saksela O., Kalkkinen N., Alitalo K.;
RT      "A novel vascular endothelial growth factor, VEGF-C, is a ligand for
RT      the Flt4 (VEGFR-3) and KDR (VEGFR-2) receptor tyrosine kinases.";
RL      EMBO J. 15:290-298(1996).
RN      [2]
RP      ERRATUM.
RX      MEDLINE=96203094; PubMed=8612600;
RA      Joukov V., Pajusola K., Kaipainen A., Chillov D., Lahtinen I., Kukk E.,
RA      Saksela O., Kalkkinen N., Alitalo K.;
RL      EMBO J. 15:1751-1751(1996).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=96312526; PubMed=8700872;
RA      Lee J., Gray A., Yuan J., Luo S.-M., Avraham H., Wood W.I.;
RT      "Vascular endothelial growth factor-related protein: a ligand and
RT      specific activator of the tyrosine kinase receptor Flt4.";
RL      Proc. Natl. Acad. Sci. U.S.A. 93:1988-1992(1996).
RN      [4]
RP      SEQUENCE FROM N.A.
RA      Fitz L., Morris J.C., Towler P.S., Long A.J., Greco R.,
RA      Burgess P., Gianotti J., Charette A., Hennessey D., Kovacic S.,
RA      Fitzgerald M., Scatireto H., Welch N., Neben S., Flimerty H.,
RA      Zollner R., Wang J., Nickbarg E., Gassaway R., Turner K.,
RA      Wood C.R.;
RL      Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.
CC      -!- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
CC      CELL GROWTH.
CC      -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC      -!- PTM: PROBABLY PROTEOLYTICALLY PROCESSED. IN THE C-TERMINUS.
CC      -!- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce/

```

```

CC or send an email to license@lsb.slb.ch).
CC EMBL: X94216; CA63907.1; -
DR EMBL: U43142; AAA85214.1; -
DR EMBL: U58111; AAB02909.1; -
DR HSSP: P15692; 1VPE.
DR MIM: 601528; -.
DR InterPro: IPR000072; -.
DR InterPro: IPR002400; -.
DR Pfam: PF00341; PDGF_1.
DR PRINTS: PR00438; GFCSKNOT.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS0278; PDGF_2; 1.
KW Mitogen; Growth factor; Glycoprotein; Signal; Repeat.
FT SIGNAL 1 2
FT PROPEP 1 102
FT CHAIN 103 419
FT DOMAIN 275 365
FT REPEAT 275 298
FT REPEAT 299 332
FT REPEAT 333 346
FT REPEAT 347 365
FT CARBOHYD 175 175
FT CARBOHYD 205 205
FT CARBOHYD 240 240
SQ SEQUENCE 419 AA; 46883 MW; 9F598719DB3E014F CRC64;

Query Match 100.0%; Score 2336; DB 1; Length 419;
Best Local Similarity 100.0%; Pred. No. 7.2e-174;
Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MHLLGFVSACSLLAALLPGPREAPAAAAAFESGLDLSDAEPDAGEAFAVASKDLEOL 60
DB 1 MHLLGFVSACSLLAALLPGPREAPAAAAAFESGLDLSDAEPDAGEAFAVASKDLEOL 60
OY 61 RSVSSVDELMATVLYPEYKMKYKQQLRKGGWQHNRQANLNSRTETIKFAAAHYNTIELK 120
DB 61 RSVSSVDELMATVLYPEYKMKYKQQLRKGGWQHNRQANLNSRTETIKFAAAHYNTIELK 120
OY 121 SIDNEMRKTCMPREVCIDVGKEFGVATNTFFKPCVSVYRGGCCNSGLQCMNTSTSY 180
DB 121 SIDNEMRKTCMPREVCIDVGKEFGVATNTFFKPCVSVYRGGCCNSGLQCMNTSTSY 180
OY 181 LSKTLEITVPLSQGPKPYTISFANHTSCRCMSKLDVYRQVSHIIRSLPATLPOCOAN 240
DB 181 LSKTLEITVPLSQGPKPYTISFANHTSCRCMSKLDVYRQVSHIIRSLPATLPOCOAN 240
OY 241 KTCPTVYMMNNHICRCLAODEFMFSSDAGDDSTDGFHDICGPKKELDETCOCVCRAGLR 300
DB 241 KTCPTVYMMNNHICRCLAODEFMFSSDAGDDSTDGFHDICGPKKELDETCOCVCRAGLR 300
OY 301 PASCGPHKELDNSCOCCVKNKLPSCGAGNEPFDNTOCQCKRCPNNOILPCKAC 360
DB 301 PASCGPHKELDNSCOCCVKNKLPSCGAGNEPFDNTOCQCKRCPNNOILPCKAC 360
OY 361 ECTESPQKLLGKGFHNTQSCYRRPCTNRQACEPGFSYSEEVRCVPSYWKRPQMS 419
DB 361 ECTESPQKLLGKGFHNTQSCYRRPCTNRQACEPGFSYSEEVRCVPSYWKRPQMS 419
DB 361 ECTESPQKLLGKGFHNTQSCYRRPCTNRQACEPGFSYSEEVRCVPSYWKRPQMS 419

RESULT 2
VEGC_MOUSE STANDARD; PRT; 415 AA.
AC P97953:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR (VEGF-C) (FLT4 LIGAND)
DE (FLT4-L).
GN VEGFC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

CC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE=97164697; PubMed=9012504;
RA Kukk E., Lymboussaki A., Taira S., Kalpainen A., Jeltsch M.,
RA Joukov V., Altalo K.,
RT "VEGF-C receptor binding and pattern of expression with VEGFR-3
RT suggests a role in lymphatic vascular development.";
RL Development 122:3829-3837(1996).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE=97388482; PubMed=9247316;
RA Fitz L.J., Morris J.C., Towler P., Long A., Burgess P., Greco R.,
RA Wang J., Gassaway R., Nickbarg E., Kovacic S., Charleita A.,
RA Giannotti J., Finerly H., Zollner R., Beier D.R., Leak L.V.,
RA Turner K.J., Wood C.R.;
RT "Characterization of murine flt4 ligand/VEGF-C";
RL Oncogene 15:613-618(1997).
CC - FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
CC CELL GROWTH.
CC - SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC - SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.slb.ch/announce/
CC or send an email to license@lsb.slb.ch).
CC -----
CC EMBL: U73620; AAC52984.1; -
DR EMBL: U58112; AAB46707.1; -
DR HSSP: P15692; 1VPE.
DR MGD: MGI:109124; Vegfc.
DR InterPro: IPR000072; -.
DR InterPro: IPR002400; -.
DR Pfam: PF00341; PDGF_1.
DR PRINTS: PR00438; GFCSKNOT.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS0278; PDGF_2; 1.
KW Mitogen; Growth factor; Glycoprotein; Signal; Repeat.
FT SIGNAL 1 2
FT PROPEP 1 98
FT CHAIN 99 415
FT DOMAIN 271 361
FT REPEAT 271 294
FT REPEAT 295 318
FT REPEAT 319 342
FT REPEAT 343 361
FT CARBOHYD 171 171
FT CARBOHYD 201 201
FT CARBOHYD 236 236
SQ SEQUENCE 415 AA; 46471 MW; D9D3DD3CEC659D6 CRC64;

Query Match 87.7%; Score 2048; DB 1; Length 415;
Best Local Similarity 85.4%; Pred. No. 1.4e-151;
Matches 358; Conservative 28; Mismatches 29; Indels 4; Gaps 1;

OY 1 MHLLGFVSACSLLAALLPGPREAPAAAAAFESGLDLSDAEPDAGEAFAVASKDLEOL 60
DB 1 MHLLGFVSACSLLAALLPGPREAPAAAAAFESGLDLSDAEPDAGEAFAVASKDLEOL 60
OY 61 RSVSSVDELMATVLYPEYKMKYKQQLRKGGWQHNRQANLNSRTETIKFAAAHYNTIELK 120
DB 61 RSVSSVDELMATVLYPEYKMKYKQQLRKGGWQHNRQANLNSRTETIKFAAAHYNTIELK 120
OY 121 SIDNEMRKTCMPREVCIDVGKEFGVATNTFFKPCVSVYRGGCCNSGLQCMNTSTSY 180
DB 121 SIDNEMRKTCMPREVCIDVGKEFGVATNTFFKPCVSVYRGGCCNSGLQCMNTSTSY 180

```

```

Db 117 SIDNEMKTCQMPREVCDVGEKEFGAATNTFFKPCVSVYRCGGCCNSEGLQCMNTSTGY 176
QY 181 LSKTLEITVPLSQGPKPYTISFANHTSCRCMSKLDYRVSHIIRRSPLPATLPQCOAN 240
Db 177 LSTLTLEITVPLSQGPKPYTISFANHTSCRCMSKLDYRVSHIIRRSPLPATLPQCOAN 236
QY 241 KTCPTNYMNNHICRCLAOEDFMFSSDAGDSTDFHDIICGPKKLEDETCQCVCRAGLR 300
Db 237 KTCPTNYMNNHICRCLAOQODFFYFSNVEDSTNGFHDYCGPKKLEDETCQCVCRAGLR 296
QY 301 PACSGHKELDNRSCQCVCKNKLFPSQCGANREPDENTCQCVCKRCPNQPPLNPKCAC 360
Db 297 PSSCGHKELDNRSCQCVCKNKLFPSQCGANREPDENTCQCVCKRCPNQPPLNPKCAC 356
QY 361 ECESPOKCLLKGRKHHOTCSYRRPCTNRKACEPGFSYSEVRCVPSYWKRPOMS 419
Db 357 ECETNRQKCLKGRKHHOTCSYRRPCTNRKACEPGFSYSEVRCVPSYWKRPHLN 415

RESULT 3
BAR3.CHITE STANDARD; PRT: 1700 AA.
AC 003376;
DT 01-OCT-1993 (rel. 27, Created)
DT 01-OCT-1993 (rel. 27, Last sequence update)
DT 01-OCT-1994 (rel. 30, Last annotation update)
DE BALBIANI RING PROTEIN 3 PRECURSOR.
GN BR3.
OS Chironomus tentans (Midge).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomidae; Chironomidae; Chironominae; Chironomus.
OX NCBI_Taxid=7153;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=90172404; PubMed=1689777;
RA Paulsson G., Lendahl U., Galli J., Ericsson C., Wieslander L.;
RT "The Balbiani ring 3 gene in Chironomus tentans has a diverged
RT repetitive structure split by many introns.";
RL J. Mol. Biol. 211:331-349(1990).
CC -I- FUNCTION: USED BY THE LARVAE TO CONSTRUCT A SUPRAMOLECULAR
CC STRUCTURE, THE LARVAL TUBE. BALBIANI RING PROTEIN 3 COULD PLAY A
CC ROLE AS A TRANSPORT PROTEIN THAT BINDS TO OTHER PROTEINS
CC INTRACELLULARLY AND IN THE GLAND LUMEN IN ORDER TO PREVENT THESE
CC FROM FORMING WATER-INSOLUBLE FIBERS TOO EARLY.
CC -I- SUBCELLULAR LOCATION: SECRETED.
CC -I- TISSUE SPECIFICITY: SALIVARY GLAND.
CC -I- DOMAIN: HAS 82 APPROXIMATE REPEATS OF CYS-X-CYS-X-CYS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: X52263; CA36506.1;
CC DR PIR: S08167; S08167.
CC DR HSSP: P18055; 2MRB.
CC DR InterPro: IPR000853;
CC DR PRINTS: PR00876; MTNEMARODE.
CC DR Repeat: signal.
CC KM SIGNAL
CC FT CHAIN 1 1700 POTENTIAL.
CC FT BALBIANI RING PROTEIN 3.
CC FT SEQUENCE 1700 AA; 186145 MW; 342028521B0815 CRC64;

```

Query Match 11.0%; Score 257; DB 1; Length 1700;
 Best Local Similarity 22.8%; Pred. No. 3,1e-12;
 Matches 89; Conservative 47; Mismatches 145; Indels 110; Gaps 18;

```

QY 79 KMYKCOLRK----GGMQHNREQANLNSRTEETIKFAAAHYNTIELKSIDNEMKTCQMPR 134
Db 727 KTKCKCEKEMPTGGCENKMKC-----DETCDCVCPKNCIARV---WDAKICS-- 775
QY 135 EVCIDVGEKEFGAATNTFFKPCVSVYRCGG-----CCNSEGLQCMNTSTSYLSKTLF 186
Db 776 --CI-----CVNPEKCNPOVLKDTCCCGGOMVSKCAPOKFI-ENIC 815
QY 187 EITVPLSQGPKPYTISFANHTSCRCMSKLDY-----YRVSHIIRRSPLPATLPQCOA 238
Db 816 DACPFPKCKCKAPLWSDFCDCVCPNSASMKTCLSPKEMNVYTCDCDNPFR--PDC-- 871
QY 239 ANKTPTNYMNNHICRC---LAQEDFMFSSDAGDSTDFHDIICG-PKKEID----- 287
Db 872 ----CGTCKKMDKCKCGCPNAGTDC-----AGGKKFNDFTCSGCCPSGKIDCTGNFTKW 922
QY 288 -EETCCVCRAGIRPASCGRPHKELDNRSCQCVCKNKLFPSQCGANREPDENTCQCVCR- 345
Db 923 SAETCTGCG--GDVNRNCCNLKFNNDLCCQCECKNKQEMANCKSPRTWNYDCKVCYKNA 980
QY 346 -----TCPRNQPPLNPKGACGECETESPOKCLLKGRKHH 378
Db 981 DSDDCVKPQIMLDQCKCGCPASAOHTCPANKRFTLEKSCGCKSPMPSPIPGKKME 1040
QY 379 QTCSCYRRPCTNRKACEPGFSYSEVRCV 409
Db 1041 DKCVY---ECAN-VKTCESGPQWCDNCKCI 1067

RESULT 4
VEGF_HUMAN STANDARD; PRT: 215 AA.
ID P15692;
AC 01-APR-1990 (rel. 14, Created)
DT 01-APR-1990 (rel. 14, Last sequence update)
DT 15-JUL-1999 (rel. 38, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR PRECURSOR (VEGF) (VASCULAR
DE PERMEABILITY FACTOR) (VPF).
GN VEGF OR VEGFA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90069608; PubMed=2479986;
RA Leung D.W., Cachianes G., Kuang W.-J., Goeddel D.V., Ferrara N.;
RT "Vascular endothelial growth factor is a secreted angiogenic
RT mitogen.";
RL Science 246:1306-1309(1989).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90069609; PubMed=2479987;
RA Keck P.J., Hauser S.D., Krivi G., Sanzo K., Warren T., Feder J.,
RA Connolly D.T.;
RT "Vascular permeability factor, an endothelial cell mitogen related to
RT PDGF.";
RL Science 246:1309-1312(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91268072; PubMed=1711045;
RA Fischer E., Mitchell R., Hartman T., Silva M., Gospodarowicz D.,
RA Fiddes J.C., Abraham J.A.;
RT "The human gene for vascular endothelial growth factor. Multiple
RT protein forms are encoded through alternative exon splicing.";
RL J. Biol. Chem. 266:11947-11954(1991).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92231879; PubMed=1567395;
RA Weindel K., Maime D., Weich H.A.;
RT "AIDS-associated Kaposi's sarcoma cells in culture express vascular
RT endothelial growth factor.";
RL Biochem. Biophys. Res. Commun. 183:1167-1174(1992).

```

[5]
 RP PRELIMINARY SEQUENCE OF 27-36: 43-50 AND 59-81.
 RA MEDLINE-90062112; PubMed-2584205;
 RA Connolly D.T., Olander J.V., Heuvelman D., Nelson R., Monsell R.,
 RA Siegel N., Haymore B.L., Leimgruber R., Feder J.;
 RA "Human vascular permeability factor. Isolation from U937 cells.";
 RA J. Biol. Chem. 264:20017-20024(1989).
 RL [6]
 RP SEQUENCE OF 27-41.
 RA MEDLINE-93145946; PubMed-7678805;
 RA Fiedlich B.L., Jaeger B., Schoellmann C., Weindel K., Witting J.,
 RA Kocms G., Marne D., Hug H., Weich H.A.;
 RA "Synthesis and assembly of functionally active human vascular
 RA endothelial growth factor homodimers in insect cells.";
 RA Eur. J. Biochem. 211:19-26(1993).
 RL [7]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 34-135.
 RA MEDLINE-97352774; PubMed-9207067;
 RA Muller Y.A., Li B., Christinger H.W., Wells J.A., Cunningham B.C.,
 RA de Vos A.M.;
 RA "Vascular endothelial growth receptor: crystal structure and functional
 RA mapping of the kinase domain receptor binding site.";
 RA Proc. Natl. Acad. Sci. U.S.A. 94:7192-7197(1997).
 RL [8]
 RP X-RAY CRYSTALLOGRAPHY (1.93 ANGSTROMS) OF 34-135.
 RA MEDLINE-98035455; PubMed-9351807;
 RA Muller Y.A., Christinger H.W., Keyt B.A., de Vos A.M.;
 RA "The crystal structure of vascular endothelial growth factor (VEGF)
 RA refined to 1.93-A resolution: multiple copy flexibility and receptor
 RA binding.";
 RA Structure 5:1325-1338(1997).
 RL [9]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 39-134.
 RA MEDLINE-99119204; PubMed-9922142;
 RA Wisemann C., Christinger H.W., Cochran A.G., Cunningham B.C.,
 RA Faltbrother W.J., Keenan C.J., Meng G., de Vos A.M.;
 RA "Crystal structure of the complex between VEGF and a receptor-blocking
 RA peptide.";
 RA Biochemistry 37:17765-17772(1998).
 RL [10]
 RP STRUCTURE BY NMR OF 34-135.
 RA MEDLINE-97477915; PubMed-9336848;
 RA Faltbrother W.J., Champe M.A., Christinger H.W., Keyt B.A.,
 RA Starovasnik M.A.;
 RA "1H, 13C, and 15N backbone assignment and secondary structure of the
 RA receptor-binding domain of vascular endothelial growth factor.";
 RA Protein Sci. 6:2250-2260(1997).
 RL [11]
 RP STRUCTURE BY NMR OF 137-215.
 RA MEDLINE-98298440; PubMed-9634701;
 RA Faltbrother W.J., Champe M.A., Christinger H.W., Keyt B.A.,
 RA Starovasnik M.A.;
 RA "Solution structure of the heparin-binding domain of vascular
 RA endothelial growth factor.";
 RA Structure 6:637-648(1998).
 RL [12]
 RP FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
 CC CELL GROWTH. INDUCES ENDOTHELIAL PROLIFERATION AND VASCULAR
 CC PERMEABILITY.
 CC -I- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -I- SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR
 CC TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN (BY
 CC SIMILARITY).
 CC -I- ALTERNATIVE PRODUCTS: FOUR FORMS OF VEGF ARE PRODUCED BY
 CC ALTERNATIVE SPLICING OF THE SAME GENE (VEGF-121, VEGF-165,
 CC VEGF-189 AND VEGF-215).
 CC -I- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL: M32977; AAA35789.1; -
 DR EMBL: M27281; AAA36807.1; -
 DR EMBL: M63978; AAA36804.1; -
 DR EMBL: M63971; AAA36804.1; JOINED.
 DR EMBL: M63972; AAA36804.1; JOINED.
 DR EMBL: M63973; AAA36804.1; JOINED.
 DR EMBL: M63974; AAA36804.1; JOINED.
 DR EMBL: M63975; AAA36804.1; JOINED.
 DR EMBL: M63976; AAA36804.1; JOINED.
 DR EMBL: M63977; AAA36804.1; JOINED.
 DR EMBL: M63978; AAA36804.1; JOINED.
 DR PIR: A34492; A34492.
 DR PIR: A40079; A40079.
 DR PIR: A40080; A40080.
 DR PIR: A40081; A40081.
 DR PIR: A40082; A40082.
 DR PIR: A40083; A40083.
 DR PIR: A40084; A40084.
 DR PIR: A40085; A40085.
 DR PIR: A40086; A40086.
 DR PIR: A40087; A40087.
 DR PIR: A40088; A40088.
 DR PIR: A40089; A40089.
 DR PIR: A40090; A40090.
 DR PIR: A40091; A40091.
 DR PIR: A40092; A40092.
 DR PIR: A40093; A40093.
 DR PIR: A40094; A40094.
 DR PIR: A40095; A40095.
 DR PIR: A40096; A40096.
 DR PIR: A40097; A40097.
 DR PIR: A40098; A40098.
 DR PIR: A40099; A40099.
 DR PIR: A40100; A40100.
 DR PIR: A40101; A40101.
 DR PIR: A40102; A40102.
 DR PIR: A40103; A40103.
 DR PIR: A40104; A40104.
 DR PIR: A40105; A40105.
 DR PIR: A40106; A40106.
 DR PIR: A40107; A40107.
 DR PIR: A40108; A40108.
 DR PIR: A40109; A40109.
 DR PIR: A40110; A40110.
 DR PIR: A40111; A40111.
 DR PIR: A40112; A40112.
 DR PIR: A40113; A40113.
 DR PIR: A40114; A40114.
 DR PIR: A40115; A40115.
 DR PIR: A40116; A40116.
 DR PIR: A40117; A40117.
 DR PIR: A40118; A40118.
 DR PIR: A40119; A40119.
 DR PIR: A40120; A40120.
 DR PIR: A40121; A40121.
 DR PIR: A40122; A40122.
 DR PIR: A40123; A40123.
 DR PIR: A40124; A40124.
 DR PIR: A40125; A40125.
 DR PIR: A40126; A40126.
 DR PIR: A40127; A40127.
 DR PIR: A40128; A40128.
 DR PIR: A40129; A40129.
 DR PIR: A40130; A40130.
 DR PIR: A40131; A40131.
 DR PIR: A40132; A40132.
 DR PIR: A40133; A40133.
 DR PIR: A40134; A40134.
 DR PIR: A40135; A40135.
 DR PIR: A40136; A40136.
 DR PIR: A40137; A40137.
 DR PIR: A40138; A40138.
 DR PIR: A40139; A40139.
 DR PIR: A40140; A40140.
 DR PIR: A40141; A40141.
 DR PIR: A40142; A40142.
 DR PIR: A40143; A40143.
 DR PIR: A40144; A40144.
 DR PIR: A40145; A40145.
 DR PIR: A40146; A40146.
 DR PIR: A40147; A40147.
 DR PIR: A40148; A40148.
 DR PIR: A40149; A40149.
 DR PIR: A40150; A40150.
 DR PIR: A40151; A40151.
 DR PIR: A40152; A40152.
 DR PIR: A40153; A40153.
 DR PIR: A40154; A40154.
 DR PIR: A40155; A40155.
 DR PIR: A40156; A40156.
 DR PIR: A40157; A40157.
 DR PIR: A40158; A40158.
 DR PIR: A40159; A40159.
 DR PIR: A40160; A40160.
 DR PIR: A40161; A40161.
 DR PIR: A40162; A40162.
 DR PIR: A40163; A40163.
 DR PIR: A40164; A40164.
 DR PIR: A40165; A40165.
 DR PIR: A40166; A40166.
 DR PIR: A40167; A40167.
 DR PIR: A40168; A40168.
 DR PIR: A40169; A40169.
 DR PIR: A40170; A40170.
 DR PIR: A40171; A40171.
 DR PIR: A40172; A40172.
 DR PIR: A40173; A40173.
 DR PIR: A40174; A40174.
 DR PIR: A40175; A40175.
 DR PIR: A40176; A40176.
 DR PIR: A40177; A40177.
 DR PIR: A40178; A40178.
 DR PIR: A40179; A40179.
 DR PIR: A40180; A40180.
 DR PIR: A40181; A40181.
 DR PIR: A40182; A40182.
 DR PIR: A40183; A40183.
 DR PIR: A40184; A40184.
 DR PIR: A40185; A40185.
 DR PIR: A40186; A40186.
 DR PIR: A40187; A40187.
 DR PIR: A40188; A40188.
 DR PIR: A40189; A40189.
 DR PIR: A40190; A40190.
 DR PIR: A40191; A40191.
 DR PIR: A40192; A40192.
 DR PIR: A40193; A40193.
 DR PIR: A40194; A40194.
 DR PIR: A40195; A40195.
 DR PIR: A40196; A40196.
 DR PIR: A40197; A40197.
 DR PIR: A40198; A40198.
 DR PIR: A40199; A40199.
 DR PIR: A40200; A40200.
 DR PIR: A40201; A40201.
 DR PIR: A40202; A40202.
 DR PIR: A40203; A40203.
 DR PIR: A40204; A40204.
 DR PIR: A40205; A40205.
 DR PIR: A40206; A40206.
 DR PIR: A40207; A40207.
 DR PIR: A40208; A40208.
 DR PIR: A40209; A40209.
 DR PIR: A40210; A40210.
 DR PIR: A40211; A40211.
 DR PIR: A40212; A40212.
 DR PIR: A40213; A40213.
 DR PIR: A40214; A40214.
 DR PIR: A40215; A40215.
 DR PIR: A40216; A40216.
 DR PIR: A40217; A40217.
 DR PIR: A40218; A40218.
 DR PIR: A40219; A40219.
 DR PIR: A40220; A40220.
 DR PIR: A40221; A40221.
 DR PIR: A40222; A40222.
 DR PIR: A40223; A40223.
 DR PIR: A40224; A40224.
 DR PIR: A40225; A40225.
 DR PIR: A40226; A40226.
 DR PIR: A40227; A40227.
 DR PIR: A40228; A40228.
 DR PIR: A40229; A40229.
 DR PIR: A40230; A40230.
 DR PIR: A40231; A40231.
 DR PIR: A40232; A40232.
 DR PIR: A40233; A40233.
 DR PIR: A40234; A40234.
 DR PIR: A40235; A40235.
 DR PIR: A40236; A40236.
 DR PIR: A40237; A40237.
 DR PIR: A40238; A40238.
 DR PIR: A40239; A40239.
 DR PIR: A40240; A40240.
 DR PIR: A40241; A40241.
 DR PIR: A40242; A40242.
 DR PIR: A40243; A40243.
 DR PIR: A40244; A40244.
 DR PIR: A40245; A40245.
 DR PIR: A40246; A40246.
 DR PIR: A40247; A40247.
 DR PIR: A40248; A40248.
 DR PIR: A40249; A40249.
 DR PIR: A40250; A40250.
 DR PIR: A40251; A40251.
 DR PIR: A40252; A40252.
 DR PIR: A40253; A40253.
 DR PIR: A40254; A40254.
 DR PIR: A40255; A40255.
 DR PIR: A40256; A40256.
 DR PIR: A40257; A40257.
 DR PIR: A40258; A40258.
 DR PIR: A40259; A40259.
 DR PIR: A40260; A40260.
 DR PIR: A40261; A40261.
 DR PIR: A40262; A40262.
 DR PIR: A40263; A40263.
 DR PIR: A40264; A40264.
 DR PIR: A40265; A40265.
 DR PIR: A40266; A40266.
 DR PIR: A40267; A40267.
 DR PIR: A40268; A40268.
 DR PIR: A40269; A40269.
 DR PIR: A40270; A40270.
 DR PIR: A40271; A40271.
 DR PIR: A40272; A40272.
 DR PIR: A40273; A40273.
 DR PIR: A40274; A40274.
 DR PIR: A40275; A40275.
 DR PIR: A40276; A40276.
 DR PIR: A40277; A40277.
 DR PIR: A40278; A40278.
 DR PIR: A40279; A40279.
 DR PIR: A40280; A40280.
 DR PIR: A40281; A40281.
 DR PIR: A40282; A40282.
 DR PIR: A40283; A40283.
 DR PIR: A40284; A40284.
 DR PIR: A40285; A40285.
 DR PIR: A40286; A40286.
 DR PIR: A40287; A40287.
 DR PIR: A40288; A40288.
 DR PIR: A40289; A40289.
 DR PIR: A40290; A40290.
 DR PIR: A40291; A40291.
 DR PIR: A40292; A40292.
 DR PIR: A40293; A40293.
 DR PIR: A40294; A40294.
 DR PIR: A40295; A40295.
 DR PIR: A40296; A40296.
 DR PIR: A40297; A40297.
 DR PIR: A40298; A40298.
 DR PIR: A40299; A40299.
 DR PIR: A40300; A40300.
 DR PIR: A40301; A40301.
 DR PIR: A40302; A40302.
 DR PIR: A40303; A40303.
 DR PIR: A40304; A40304.
 DR PIR: A40305; A40305.
 DR PIR: A40306; A40306.
 DR PIR: A40307; A40307.
 DR PIR: A40308; A40308.
 DR PIR: A40309; A40309.
 DR PIR: A40310; A40310.
 DR PIR: A40311; A40311.
 DR PIR: A40312; A40312.
 DR PIR: A40313; A40313.
 DR PIR: A40314; A40314.
 DR PIR: A40315; A40315.
 DR PIR: A40316; A40316.
 DR PIR: A40317; A40317.
 DR PIR: A40318; A40318.
 DR PIR: A40319; A40319.
 DR PIR: A40320; A40320.
 DR PIR: A40321; A40321.
 DR PIR: A40322; A40322.
 DR PIR: A40323; A40323.
 DR PIR: A40324; A40324.
 DR PIR: A40325; A40325.
 DR PIR: A40326; A40326.
 DR PIR: A40327; A40327.
 DR PIR: A40328; A40328.
 DR PIR: A40329; A40329.
 DR PIR: A40330; A40330.
 DR PIR: A40331; A40331.
 DR PIR: A40332; A40332.
 DR PIR: A40333; A40333.
 DR PIR: A40334; A40334.
 DR PIR: A40335; A40335.
 DR PIR: A40336; A40336.
 DR PIR: A40337; A40337.
 DR PIR: A40338; A40338.
 DR PIR: A40339; A40339.
 DR PIR: A40340; A40340.
 DR PIR: A40341; A40341.
 DR PIR: A40342; A40342.
 DR PIR: A40343; A40343.
 DR PIR: A40344; A40344.
 DR PIR: A40345; A40345.
 DR PIR: A40346; A40346.
 DR PIR: A40347; A40347.
 DR PIR: A40348; A40348.
 DR PIR: A40349; A40349.
 DR PIR: A40350; A40350.
 DR PIR: A40351; A40351.
 DR PIR: A40352; A40352.
 DR PIR: A40353; A40353.
 DR PIR: A40354; A40354.
 DR PIR: A40355; A40355.
 DR PIR: A40356; A40356.
 DR PIR: A40357; A40357.
 DR PIR: A40358; A40358.
 DR PIR: A40359; A40359.
 DR PIR: A40360; A40360.
 DR PIR: A40361; A40361.
 DR PIR: A40362; A40362.
 DR PIR: A40363; A40363.
 DR PIR: A40364; A40364.
 DR PIR: A40365; A40365.
 DR PIR: A40366; A40366.
 DR PIR: A40367; A40367.
 DR PIR: A40368; A40368.
 DR PIR: A40369; A40369.
 DR PIR: A40370; A40370.
 DR PIR: A40371; A40371.
 DR PIR: A40372; A40372.
 DR PIR: A40373; A40373.
 DR PIR: A40374; A40374.
 DR PIR: A40375; A40375.
 DR PIR: A40376; A40376.
 DR PIR: A40377; A40377.
 DR PIR: A40378; A40378.
 DR PIR: A40379; A40379.
 DR PIR: A40380; A40380.
 DR PIR: A40381; A40381.
 DR PIR: A40382; A40382.
 DR PIR: A40383; A40383.
 DR PIR: A40384; A40384.
 DR PIR: A40385; A40385.
 DR PIR: A40386; A40386.
 DR PIR: A40387; A40387.
 DR PIR: A40388; A40388.
 DR PIR: A40389; A40389.
 DR PIR: A40390; A40390.
 DR PIR: A40391; A40391.
 DR PIR: A40392; A40392.
 DR PIR: A40393; A40393.
 DR PIR: A40394; A40394.
 DR PIR: A40395; A40395.
 DR PIR: A40396; A40396.
 DR PIR: A40397; A40397.
 DR PIR: A40398; A40398.
 DR PIR: A40399; A40399.
 DR PIR: A40400; A40400.
 DR PIR: A40401; A40401.
 DR PIR: A40402; A40402.
 DR PIR: A40403; A40403.
 DR PIR: A40404; A40404.
 DR PIR: A40405; A40405.
 DR PIR: A40406; A40406.
 DR PIR: A40407; A40407.
 DR PIR: A40408; A40408.
 DR PIR: A40409; A40409.
 DR PIR: A40410; A40410.
 DR PIR: A40411; A40411.
 DR PIR: A40412; A40412.
 DR PIR: A40413; A40413.
 DR PIR: A40414; A40414.
 DR PIR: A40415; A40415.
 DR PIR: A40416; A40416.
 DR PIR: A40417; A40417.
 DR PIR: A40418; A40418.
 DR PIR: A40419; A40419.
 DR PIR: A40420; A40420.
 DR PIR: A40421; A40421.
 DR PIR: A40422; A40422.
 DR PIR: A40423; A40423.
 DR PIR: A40424; A40424.
 DR PIR: A40425; A40425.
 DR PIR: A40426; A40426.
 DR PIR: A40427; A40427.
 DR PIR: A40428; A40428.
 DR PIR: A40429; A40429.
 DR PIR: A40430; A40430.
 DR PIR: A40431; A40431.
 DR PIR: A40432; A40432.
 DR PIR: A40433; A40433.
 DR PIR: A40434; A40434.
 DR PIR: A40435; A40435.
 DR PIR: A40436; A40436.
 DR PIR: A40437; A40437.
 DR PIR: A40438; A40438.
 DR PIR: A40439; A40439.
 DR PIR: A40440; A40440.
 DR PIR: A40441; A40441.
 DR PIR: A40442; A40442.
 DR PIR: A40443; A40443.
 DR PIR: A40444; A40444.
 DR PIR: A40445; A40445.
 DR PIR: A40446; A40446.
 DR PIR: A40447; A40447.
 DR PIR: A40448; A40448.
 DR PIR: A40449; A40449.
 DR PIR: A40450; A40450.
 DR PIR: A40451; A40451.
 DR PIR: A40452; A40452.
 DR PIR: A40453; A40453.
 DR PIR: A40454; A40454.
 DR PIR: A40455; A40455.
 DR PIR: A40456; A40456.
 DR PIR: A40457; A40457.
 DR PIR: A40458; A40458.
 DR PIR: A40459; A40459.
 DR PIR: A40460; A40460.
 DR PIR: A40461; A40461.
 DR PIR: A40462; A40462.
 DR PIR: A40463; A40463.
 DR PIR: A40464; A40464.
 DR PIR: A40465; A40465.
 DR PIR: A40466; A40466.
 DR PIR: A40467; A40467.
 DR PIR: A40468; A40468.
 DR PIR: A40469; A40469.
 DR PIR: A40470; A40470.
 DR PIR: A40471; A40471.
 DR PIR: A40472; A40472.
 DR PIR: A40473; A40473.
 DR PIR: A40474; A40474.
 DR PIR: A40475; A40475.
 DR PIR: A40476; A40476.
 DR PIR: A40477; A40477.
 DR PIR: A40478; A40478.
 DR PIR: A40479; A40479.
 DR PIR: A40480; A40480.
 DR PIR: A40481; A40481.
 DR PIR: A40482; A40482.
 DR PIR: A40483; A40483.
 DR PIR: A40484; A40484.
 DR PIR: A40485; A40485.
 DR PIR: A40486; A40486.
 DR PIR: A40487; A40487.
 DR PIR: A40488; A40488.
 DR PIR: A40489; A40489.
 DR PIR: A40490; A40490.
 DR PIR: A40491; A40491.
 DR PIR: A40492; A40492.
 DR PIR: A40493; A40493.
 DR PIR: A40494; A40494.
 DR PIR: A40495; A40495.
 DR PIR: A40496; A40496.
 DR PIR: A40497; A40497.
 DR PIR: A40498; A40498.
 DR PIR: A40499; A40499.
 DR PIR: A40500; A40500.
 DR PIR: A40501; A40501.
 DR PIR: A40502; A40502.
 DR PIR: A40503; A40503.
 DR PIR: A40504; A40504.
 DR PIR: A40505; A40505.
 DR PIR: A40506; A40506.
 DR PIR: A40507; A40507.
 DR PIR: A40508; A40508.
 DR PIR: A40509; A40509.
 DR PIR: A40510; A40510.
 DR PIR: A40511; A40511.
 DR PIR: A40512; A40512.
 DR PIR: A40513; A40513.
 DR PIR: A40514; A40514.
 DR PIR: A40515; A40515.
 DR PIR: A40516; A40516.
 DR PIR: A40517; A40517.
 DR PIR: A40518; A40518.
 DR PIR: A40519; A40519.
 DR PIR: A40520; A40520.
 DR PIR: A40521; A40521.
 DR PIR: A40522; A40522.
 DR PIR: A40523; A40523.
 DR PIR: A40524; A40524.
 DR PIR: A40525; A40525.
 DR PIR: A40526; A40526.
 DR PIR: A40527; A40527.
 DR PIR: A40528; A40528.
 DR PIR: A40529; A40529.
 DR PIR: A40530; A40530.
 DR PIR: A40531; A40531.
 DR PIR: A40532; A40532.
 DR PIR: A40533; A40533.
 DR PIR: A40534; A40534.
 DR PIR: A40535; A40535.
 DR PIR: A40536; A40536.
 DR PIR: A40537; A40537.
 DR PIR: A40538; A40538.
 DR PIR: A40539; A40539.
 DR PIR: A40540; A40540.
 DR PIR: A40541; A40541.
 DR PIR: A40542; A40542.
 DR PIR: A40543; A40543.
 DR PIR: A40544; A40544.


```

RESULT 5
VEGF_PIG
ID VEGF_PIG STANDARD: PRT: 190 AA.
AC P49151:
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR PRECURSOR (VEGF) (VASCULAR
DE PERMEABILITY FACTOR) (VPF).
GN VEGF.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RA MEDLINE=95143284; PubMed=7841203;
RA Sharma H.S., Tang Z.H., Cho B.C.H., Verdouw P.D.;
RT "Nucleotide sequence and expression of the porcine vascular
RT endothelial growth factor";
RL Biochim. Biophys. Acta 1260:235-238(1995).
CC -1- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
CC CELL GROWTH. INDUCES ENDOTHELIAL PROLIFERATION AND VASCULAR
CC PERMEABILITY (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR
CC TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X81380; CAA57143.1; -
DR HSPF: P15692; 2VGH.
DR InterPro: IPR000072; -
DR Pfam: PF00341; PDGF_1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS00278; PDGF_2; 1.
KW Mitogen; Growth factor; Glycoprotein; Signal.
FT SIGNAL 1 26
FT CHAIN 27 190
FT DISULFID 51 93
FT DISULFID 82 127
FT DISULFID 86 129
FT DISULFID 76 76
FT DISULFID 85 85
FT CARBOHYD 100 100
SQ SEQUENCE 190 AA: 22368 MW: 04D4DB8D913047F CRC64;
-----
Query Match 9.4%; Score 218.5; DB 1; Length 190;
Best Local Similarity 23.3%; Pred. No. 3.1e-10;
Matches 62; Conservative 25; Mismatches 72; Indels 107; Gaps 8;

```

```

OY 260 EDFMFSADGDDSTDFHICGPNKELDETCQCYCRAGLRPASCGRHEL---DRNSQ 316
DB 139 -----ENPCGPCSE-----RRKHLFVQDPQYCK 161
OY 317 CVCKNRKLPSPSGCANR-EFEDNTCC 341
DB 162 CSCKNT--DSRCKARQLELNETRC 185
-----
RESULT 6
VEGF_MOUSE
ID VEGF_MOUSE STANDARD: PRT: 214 AA.
AC 000731:
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR PRECURSOR (VEGF) (VASCULAR
DE PERMEABILITY FACTOR) (VPF).
GN VEGF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=92274860; PubMed=1592003;
RA Breier G., Albrecht U., Steier S., Risau W.;
RT "Expression of vascular endothelial growth factor during embryonic
RT angiogenesis and endothelial cell differentiation.";
RL Development 114:521-532(1992).
[2]
RP SEQUENCE FROM N.A. (VEGF-1).
RA MEDLINE=92355593; PubMed=1644816;
RA Claffey K.P., Wilkison W.O., Spiegelman B.M.;
RT "Vascular endothelial growth factor. Regulation by cell
RT differentiation and activated second messenger pathways.";
RL J. Biol. Chem. 267:16317-16322(1992).
[3]
RP SEQUENCE OF 1-3 FROM N.A.
RA MEDLINE=96216498; PubMed=8632007;
RA Shima D.T., Kuroki M., Deutsch U., Ng Y., Adams A.P., D'Amore P.A.;
RT "The mouse gene for vascular endothelial growth factor. Genomic
RT structure, definition of the transcriptional unit, and
RT characterization of transcriptional and post-transcriptional
RT regulatory sequences.";
RL J. Biol. Chem. 271:3877-3883(1996).
CC -1- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
CC CELL GROWTH. INDUCES ENDOTHELIAL PROLIFERATION AND VASCULAR
CC PERMEABILITY.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: VEGF-1 AND VEGF-2 ARE SECRETED WHILE
CC VEGF-3 REMAINS CELL-SURFACE ASSOCIATED UNLESS RELEASED BY
CC HEPARIN.
CC -1- ALTERNATIVE PRODUCTS: THREE FORMS (VEGF-1, VEGF-2 AND VEGF-3) ARE
CC PRODUCED AS A RESULT OF ALTERNATIVE SPLICING OF THE SAME GENE. THE
CC LONGEST FORM (VEGF-3, SHOWN HERE) CONTAINS A BASIC INSERT LINKED
CC TO CELL-ASSOCIATION/HEPARIN-BINDING.
CC -1- TISSUE SPECIFICITY: IN DEVELOPING EMBRYOS, EXPRESSED MAINLY IN
CC THE CHOROID PLEXUS, PARAVENTRICULAR NEUROEPITHELIUM, PLACENTA AND
CC KIDNEY GLOMERULI. ALSO FOUND IN BRONCHIAL EPITHELIUM, ADRENAL
CC GLAND AND IN SEMINIFEROUS TUBULES OF TESTIS. HIGH EXPRESSION OF
CC VEGF CONTINUES IN KIDNEY GLOMERULI AND CHOROID PLEXUS IN ADULTS.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----

```

```

DR EMBL: S37052; AAB2252.1; -
DR EMBL: S38083; AAB2253.1; -
DR EMBL: S38100; AAB2254.1; -
DR EMBL: M95200; AAA40547.1; -
DR EMBL: U41383; -, NOT_ANNOTATED_CDS.
DR PIR: A43351; A43351.
DR HSSP: P15692; 2VGH.
DR MGD: MGI:103178; Vegf.
DR InterPro: IPR000072; -.
DR Pfam: PF00341; PDGF_1; 1.
DR PROSITE: PS50278; PDGF_2; 1.
DR PROSITE: PS50249; PDGF_1; 1.
KW Mitogen; Growth factor; Glycoprotein; Alternative splicing; Signal.
FT SIGNAL 1 26
FT CHAIN 1 214
FT DISULFID 51 93
FT DISULFID 82 127
FT DISULFID 86 129
FT DISULFID 76 76
FT DISULFID 85 85
FT CARBOHYD 100 100
FT CARBOHYD 140 140
FT VARSPPLIC 141 164
FT VARSPPLIC 141 164
FT VARSPPLIC 141 164
FT CONFLICT 117 118
SQ SEQUENCE 214 AA; 25283 MW; B5540B51E4B6E17 CRC64;

```

```

Query Match 8.9%; Score 207.5; DB 1; Length 214;
Best Local Similarity 22.7%; Pred. No. 2.5e-09;
Matches 59; Conservative 32; Mismatches 98; Indels 71; Gaps 7;

```

```

QY 85 LRRGQWNRREOANLNSRTEETIKFAAHYNTIELKSIDNEMRKTCQMPREVCIQDGRKEF 144
DB 18 LHHAKWSQAAPRTTEGQKSHVYIKFM-----DVYRSTCRPIETLVDFQDEY 64
QY 145 GVAATNFFRPVSVYRGGCCNSEGLQCMNTSTYLSKTLPEITVPLSQ--GPKPVITIS 202
DB 65 PDEIETIFRPSCVPLMRKAGCCNDLEECVPTSESNITQIMKIPHOSQHG---EMS 120
QY 203 FANHTSCRCMSKIDYRQVHSIIRSLPATLPQCOAANKTCPTNYMNNHICRLAQDEDF 262
DB 121 FLOHSHCECRPKKDRTPKPKSVRGK---GKGQRRKRRKSRFKSMVSHCEPSERRKH 175
QY 263 MFSDDGSDSTGFHDICGPNKELDEETCCVCVCRAGLRPASCGRPHKELDRNSCQVCYCKK 322
DB 176 LRVQDP-----QTKCKSCKNT 191
QY 323 LFPSCGQANR-EFDENTCCQ 341
DB 192 --DSRCKARQLELNEKTRC 209

```

RESULT 7

```

VEGF_RAT STANDARD; PRT; 190 AA.
AC P16612;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR PRECURSOR (VEGF) (VASCULAR PERMEABILITY FACTOR) (VPF).
DE VEGF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-190.
RX MEDLINE=90207249; PubMed=2320579;
RA Conn G., Bayne M.L., Soderman D.D., Kwok P.W., Sullivan K.A.,
RA Pallis T.M., Hope D.A., Thomas K.A.;
RT "Amino acid and cDNA sequences of a vascular endothelial cell mitogen

```

```

RT that is homologous to platelet-derived growth factor.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:2628-2633(1990).
CC -! FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL CELL GROWTH. INDUCES ENDOTHELIAL PROLIFERATION AND VASCULAR PERMEABILITY.
CC -! SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -! SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN (BY SIMILARITY).
CC -! TISSUE SPECIFICITY: EXPRESSED IN THE PITUITARY, IN BRAIN, IN PARTICULARLY IN SUPRAOPTIC AND PARAVENTRICULAR NUCLEI AND THE CHOROID PLEXUS. ALSO FOUND ABUNDANTLY IN THE CORPUS LUTEUM OF THE OVARY AND IN KIDNEY GLOMERULI.
CC -! SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC
DR EMBL: M32167; AAA41211.1; -
DR PIR: A35987; A35987.
DR HSSP: P15692; 2VGH.
DR InterPro: IPR000072; -.
DR Pfam: PF00341; PDGF_1; 1.
DR PROSITE: PS50249; PDGF_1; 1.
DR PROSITE: PS50278; PDGF_2; 1.
KW Mitogen; Growth factor; Glycoprotein; Signal.
FT SIGNAL 1 26
FT CHAIN 27 190
FT DISULFID 51 93
FT DISULFID 82 127
FT DISULFID 86 129
FT DISULFID 76 76
FT DISULFID 85 85
FT CARBOHYD 100 100
SQ SEQUENCE 190 AA; 22396 MW; 589374010441F377 CRC64;

```

```

Query Match 8.8%; Score 206.5; DB 1; Length 190;
Best Local Similarity 24.5%; Pred. No. 2.6e-09;
Matches 62; Conservative 31; Mismatches 71; Indels 89; Gaps 10;

```

```

QY 92 HNRDOANLNSRTEETIKFAAHYNTIELKSIDNEMRKTCQMPREVCIQDGRKEFGVATNTF 151
DB 19 LHHAKWSQAAPRTTEGQK--AH---EVYKFM-DVYRSTCRPIETLVDFQDEYIPDEIETI 71
QY 152 FRPCVSVYRGGCCNSEGLQCMNTSTYLSKTLPEITVPLSQ--GPKPVITISFANHTSC 209
DB 72 FRPCVPLMRKAGCCNDLEECVPTSESNITQIMKIPHOSQHG---EMSFLOHSHC 127
QY 210 RCMKSLDYRQVHSIIRSLPATLPQCOAANKTCPTNYMNNHICRLAQDEDFMFSDDAG 269
DB 128 ECRPKKD-----RTKP-----ENHCRPSEGRKRNHLFVQDP- 157
QY 270 DDSTGFDHICGPNKELDEETCCVCVCRAGLRPASCGRPHKELDRNSCQVCYCKKLPSCG 329
DB 158 -----QTKCKSCKNT--DSRCK 172
QY 330 ANR-EFDENTCCQ 341
DB 173 ARQLELNEKTRC 185

```

RESULT 8

```

VEGF_BOVIN STANDARD; PRT; 190 AA.
ID VEGF_BOVIN
AC P15691;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)

```

01-OCT-1996 (Rel. 34, last annotation update)
 VASCULAR ENDOTHELIAL GROWTH FACTOR PRECURSOR (VEGF) (VASCULAR
 DE PERMEABILITY FACTOR) (VPF).
 GN VEGF.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovinae; Bovinae; Bos.
 NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-47.
 RA MEDLINE=90069608; PubMed=2479986;
 RX Leung D.W., Cachianes G., Kiang W.-J., Goeddel D.V., Ferrara N.;
 RT "Vascular endothelial growth factor is a secreted angiogenic
 mitogen.";
 RL Science 246:1306-1309(1989).
 RN [2]
 RP SEQUENCE OF 27-190 FROM N.A.
 RA MEDLINE=90121225; PubMed=2610687;
 RX Fischer E., Gospodarowicz D., Mitchell R., Silva M., Schilling J.,
 RA Lau K., Crisp T., Fiddes J.C., Abraham J.A.;
 RT "Vascular endothelial growth factor: a new member of the platelet-
 derived growth factor gene family.";
 RL Biochem. Biophys. Res. Commun. 165:1198-1206(1989).
 RN [3]
 RP SEQUENCE OF 27-31.
 RX MEDLINE=89286596; PubMed=2735925;
 RA Ferrara N., Henzel W.J.;
 RT "Placental follicular cells secrete a novel heparin-binding growth
 factor specific for vascular endothelial cells.";
 RL Biochem. Biophys. Res. Commun. 161:851-858(1989).
 CC -1- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
 CELL GROWTH. INDUCES ENDOTHELIAL PROLIFERATION AND VASCULAR
 PERMEABILITY.
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -1- SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR
 TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN (BY
 SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 CC EMBL: M32976; AAA30502.1;
 DR EMBL: M31836; AAA30804.1;
 DR EMBL: M33750; AAA30805.1;
 DR PIR: A33255; A33255.
 DR PIR: A33787; A33787.
 DR PIR: B40080; B40080.
 DR HSP: P15692; 2VGH.
 DR InterPro: IPR000072;
 DR Pfam: PF00341; PDGF_1;
 DR PROSITE: PS00249; PDGF_1;
 DR PROSITE: PS50278; PDGF_2;
 DR Mitogen: Growth factor; Glycoprotein; Alternative splicing; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 190 VASCULAR ENDOTHELIAL GROWTH FACTOR.
 FT DISULFID 51 93 BY SIMILARITY.
 FT DISULFID 82 127 BY SIMILARITY.
 FT DISULFID 86 129 BY SIMILARITY.
 FT DISULFID 76 76 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 85 85 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 139 183 MISSING (IN ISOFORM BETA).
 FT VARSPLIC 184 184 R -> K (IN ISOFORM BETA).
 SO SEQUENCE 190 AA; 22310 MW; EDBF903E46E24789 CRC64;

Query Match 8.8%; Score 205.5; DB 1; Length 190;
 Best Local Similarity 22.8%; Pred No. 3.1e-09;
 Matches 60; Conservative 25; Mismatches 77; Indels 101; Gaps 8;
 QY 85 LRRGQOHNRQGANLNSRTETIKFAAHYNTILKSIDNMRKTCQMPREYCIDYGRKF 144
 DB 18 LHHAKWSQAAPMAEGQKHEVYKFM-----DYQSRSCRPETLVDFQY 64
 QY 145 GVAATNFFRPVSVYRCGCCNSEGLQCMNTSTYLSLTLEIYPLSQ--GPKVYIS 202
 DB 65 PDEIEFIFRPSVPLMRCCGCCNDESLCEVPTEEFNITQIMIKRHSQHTG---EMS 120
 QY 203 FANHTSCRCMSKLDVROYHSIIRSLPATLPQCAANTCPTNYWMNNHICRLAQEDF 262
 DB 121 FLOHMKCECRPKKDKARQ----- 138
 QY 263 MFSSDAGDSDTGFDHICGPNKELDEETQCYCAGLRPASCGRKEL--DRNSQCVC 319
 DB 139 -----ENPCGPSE-----RRKHLFPQDPQTCCKSC 164
 QY 320 KKKLFPSCGANR-EPDENTCOC 341
 DB 165 KMT--DSRCKARQLELNEFTCRC 185
 RESULT 9
 VEGF_CAVPO STANDARD: PRT; 164 AA.
 ID VEGF_CAVPO
 AC P26617;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR (VEGF) (VASCULAR PERMEABILITY
 FACTOR) (VPF).
 GN VEGF.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Hystriognathi; Cavidae; Cavia.
 NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Berse B.;
 RT Submitted (xxx-1992) to the EMBL/Genbank/DBJ databases.
 RL -1- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
 CELL GROWTH. INDUCES ENDOTHELIAL PROLIFERATION AND VASCULAR
 PERMEABILITY.
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -1- SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR
 TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN (BY
 SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 CC EMBL: M84230; AAA37057.1;
 DR HSP: P15692; 2VGH.
 DR InterPro: IPR000072;
 DR Pfam: PF00341; PDGF_1;
 DR PROSITE: PS00249; PDGF_1;
 DR PROSITE: PS50278; PDGF_2;
 DR Mitogen: Growth factor; Glycoprotein.
 FT DISULFID 25 67 BY SIMILARITY.
 FT DISULFID 56 101 BY SIMILARITY.
 FT DISULFID 60 103 BY SIMILARITY.
 FT DISULFID 50 50 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 59 59 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).


```

OY 256 CL-----AODEPMF-----SSDAGDSTGDFHD-----ICGPKKELDE 288
DB 162 CVDGTSGLPCQGNACKPNCPSNGGTCSDADKEDYSCCECRPGYFPECERYVAPNCPKANG 241
OY 289 ETC-----QCVCRAGLRPASC-----GPKKELDR-----NSCOCVCKNKLFPSC 328
DB 242 GICSSDGSQGYRCRCRGKGGVCKVYCKPCKNKGSCVCKNGSSYNCKGKGYSGPTC 301
OY 329 GAN-----REPENT-----CCCV-----CK-----RTCPRPQPLNPGC----- 358
DB 302 GBNVCKPNCPCNRCGRCPDPSDDGFCRCVGYKGPCTCEDKPNPENTKPKCKNGKCNVNG 361
OY 359 ---ACECT-----ESPCKCLKGKRF-----HQTG 381
DB 362 KTYTCKCAVGMGRHCTDKATYKPNPCVYKPKCKNKGKCLMNGKAIKRCAYGSGRHCT 421
OY 382 SCYR-----RPTNRORACEPFGFSYSEVCRCPVSY 412
DB 422 KSYKKNPCASRCKNRKCTDKNGY---VCKCARGY 455

RESULT 11
VEGF_CHICK STANDARD: PRT: 216 AA.
ID VEGF_CHICK STANDARD: PRT: 216 AA.
AC P52582; Q91420;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR PRECURSOR (VEGF) (VASCULAR PERMEABILITY FACTOR) (VPF).
DE VEGF.
OS Gallus gallus (Chicken), and
OS Coturnix coturnix Japonica (Japanese quail)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031, 93934;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Chicken; TISSUE=Heart;
RA Takahashi T.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=C. Japonica; TISSUE=Embryo;
RX MEDLINE=96005007; PubMed=7556923;
RA Flame I., von Reutern M., Drexler H.C., Syed-Ali S., Risau W.;
RT "Overexpression of vascular endothelial growth factor in the avian embryo induces hypervascularization and increased vascular permeability without alterations of embryonic pattern formation.";
RT Dev. Biol. 171:399-414(1995).
RN [3]
RP SEQUENCE OF 60-187 FROM N.A.
RC SPECIES=C. Japonica;
RX MEDLINE=95301109; PubMed=7781909;
RA Flame I., Breier G., Risau W.;
RT "Vascular endothelial growth factor (VEGF) and VEGF receptor 2 (Flk-1) are expressed during vasculogenesis and vascular differentiation in the quail embryo.";
RT Dev. Biol. 169:699-712(1995).
RN [4]
RP FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL CELL GROWTH. INDUCES ENDOTHELIAL PROLIFERATION AND VASCULAR PERMEABILITY.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: THREE ISOFORMS (VEGF-190, VEGF-146 AND VEGF-166) ARE PRODUCED AS A RESULT OF ALTERNATIVE SPLICING OF THE SAME GENE. THE LONGER FORM CONTAINS A BASIC INSERT WHICH ACTS AS A CELL RETENTION SIGNAL.
CC -1- TISSUE SPECIFICITY: ABUNDANTLY AND EQUALLY EXPRESSED IN HEART AND LIVER, IN KIDNEY GLOMERULI, BRAIN AND YOLK SAC, VEGF-166 FORM IS 5- TO 10- TIMES MORE ABUNDANT THAN THE VEGF-190 FORM.
CC -1- DEVELOPMENTAL STAGE: THE VEGF-166 FORM IS EXPRESSED EARLY AT DAY 1

```

```

CC AND IS UPGRADED DURING GASTRULATION. EXPRESSION OF THE VEGF-190
CC FORM IS DETECTABLE ONLY FROM DAY 2.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as their content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC
CC EMBL: AB011078; BAA24925.1;
CC DR EMBL: S79680; AAB35371.1;
CC DR HSSP: P15692; 2VGH
CC DR InterPro: IPR000072;
CC DR Pfam: PF00341; PDGF_1;
CC DR PROSITE: PS00249; PDGF_1;
CC DR PROSITE: PS0278; PDGF_2;
CC KW Mitogen; Growth factor; Glycoprotein; Alternative splicing; Signal.
CC SIGNAL 1 26 BY SIMILARITY.
CC FT CHAIN 27 216 VASCULAR ENDOTHELIAL GROWTH FACTOR.
CC FT DISULFID 83 128 BY SIMILARITY.
CC FT DISULFID 87 130 BY SIMILARITY.
CC FT DISULFID 77 77 INTERCHAIN (BY SIMILARITY).
CC FT CARBOHYD 86 86 INTERCHAIN (BY SIMILARITY).
CC FT VARSPLIC 101 101 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT VARSPLIC 142 142 K->N (IN ISOFORM VEGF-166).
CC FT VARSPLIC 143 166 MISSING (IN ISOFORM VEGF-166).
CC FT VARSPLIC 166 166 F->L (IN ISOFORM VEGF-146).
CC FT VARSPLIC 167 210 MISSING (IN ISOFORM VEGF-146).
CC SQ SEQUENCE 216 AA; 25203 MW; 82E69C2F6F6GDAV CRC64;

```

```

Query Match 7.6%; Score 178.5; DB 1; Length 216;
Best Local Similarity 20.9%; Pred. No. 4.4e-07;
Matches 55; Conservative 41; Mismatches 98; Indels 69; Gaps 8;
OY 90 WOHNREOANLNSRTEETIKFAA-----HYNTEILKSDENRKTQCPREVCIOVKEF 144
DB 7 WTHMGLALLYQSAELSAAPALDGERKPNVTKFE-VERSFCRTIEVLVDFEY 65
OY 145 GVAATFPKPPCVSVYRCGCCNSEGLQCMNTSTYLSKTEITVPLS-----OGPRPV 199
DB 66 PDEVEYIIRPSCVPLMRCAGCGCGDEGLCVPD-----YVNMELARIKPHOSQHA 118
OY 200 TISFANHTSCRCMSKLDVYROYHSIIRSLPRLPQCAANKTCPTNYMNNHICRLAQ 259
DB 119 HMFLOHSEKCDPRPKDYKNNKQEKSKRGKGRKKRKKRKKRKKRKKRKKRKKRKKR 174
OY 260 EDFMSSDAGDSTGDFHDICGPKKELDEBETQCVCRAGLRASCGRPKKELDRNSCQVC 319
DB 175 RKHLFVQ-----DPQTCCKSC----- 190
OY 320 KNLFPSCGANR-EFDENTCC 341
DB 191 -KFLDSCKSKQLELNERTCRC 211

RESULT 12
VEGF_HUMAN STANDARD: PRT: 188 AA.
ID VEGF_HUMAN STANDARD: PRT: 188 AA.
AC P49765;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR B PRECURSOR (VEGF-B) (VEGF RELATED FACTOR).
GN VEGFB OR VRF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-96197355; PubMed-8637916;
RA Olofsson B., Pajusola K., Kaipainen A., von Euler G., Joukov V.,
RA Sakela O., Orpana A., Petersson R.F., Alltalo K., Eriksson U.;
RT "Vascular endothelial growth factor B, a novel growth factor for
RT endothelial cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:2576-2581(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-97077124; PubMed-8919691;
RA Grimmond S., Lagercrantz J., Drinkwater C., Sillins G., Townson S.,
RA Pollock P., Gotley D., Carson E., Rakar S., Nordenskjöld M., Ward L.,
RA Hayward N., Weber G.;
RT "Cloning and characterization of a novel human gene related to
RT vascular endothelial growth factor.";
RL Genome Res. 6:124-131(1996).
CC -1- FUNCTION: GROWTH FACTOR FOR ENDOTHELIAL CELLS. BINDS HEPARIN.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED. CAN ALSO FORM HETERODIMER
CC WITH VEGF.
CC -1- SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR
CC TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXCEPT LIVER.
CC HIGHEST LEVELS FOUND IN HEART, SKELETAL MUSCLE AND PANCREAS.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U48801; AAB06274.1; -
DR EMBL: U43369; AAA91463.1; -
DR HSSP: P15692; IYPF.
DR MIM: 601398; -
DR InterPro: IPR000072; -
DR Pfam: PF00341; PDGF_1;
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS00278; PDGF_2; 1.
KM Mitogen; Growth factor; Signal; Heparin-binding.
FT SIGNAL 1 21
FT CHAIN 22 188
FT SEQUENCE 188 AA; 21261 MW; F04654D5A3727194 CRC64;
SQ
Query Match 7.5%; Score 175.5; DB 1; Length 188;
Best Local Similarity 22.5%; Pred. No. 6.5e-07;
Matches 52; Conservative 26; Mismatches 78; Indels 75; Gaps 6;
OY 111 AAHYMTIELKSIDNEMRKQCPREVCIDVGEKFGVATNTFFKPCVSVYRCGGCCNSG 170
DB 28 AFGHQRKVVSWIDVYTRAT-CQPREVVVPLIYELMGYAKOLVPCSVYQRCGGCCPPDG 86
OY 171 LQCMNTSTSYLSKTLFEITVPLSGPKPVYISFANHSCRCMSKLDVYRQVHSIIRSLP 230
DB 87 LECVPTGHOVMQILMIRYPSQLGE---MSLEHSHQCECRPK-----KDSAVKPPDSP 138
OY 231 ATLPCQQAANKCTPTNYMNNHICGLAODEPMFSSDAGDDSTDFGHIDICGNKLEDEET 290
DB 139 RPL-----CP-----RCTOH----- 148
OY 291 CQCVCRAGLRPASGPHKELDNSCCVCKNKLFPSCGANCPEDEMTCC 341
DB 149 -----HQRPPRTCRRCRRRSFLRCQGRGLNPDTGCR 183
RESULT 13
VEGB_MOUSE STANDARD: PRT: 188 AA.
AC P49766;

```

```

DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-DEC-1998 (Rel. 37, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR B PRECURSOR (VEGF-B) (VASCULAR
DE ENDOTHELIAL GROWTH FACTOR RELATED PROTEIN) (VRF).
GN VEGFB OR VRF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-96197355; PubMed-8637916;
RA Olofsson B., Pajusola K., Kaipainen A., von Euler G., Joukov V.,
RA Sakela O., Orpana A., Petersson R.F., Alltalo K., Eriksson U.;
RT "Vascular endothelial growth factor B, a novel growth factor for
RT endothelial cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:2576-2581(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-96183052; PubMed-8607868;
RA Townson S., Lagercrantz J., Grimmond S., Sillins G.,
RA Nordenskjöld M., Weber G., Hayward N.K.;
RT "Characterization of the murine VEGF-related factor gene.";
RL Biochem. Biophys. Res. Commun. 220:922-928(1996).
CC -1- FUNCTION: GROWTH FACTOR FOR ENDOTHELIAL CELLS. BINDS HEPARIN.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED. CAN ALSO FORM HETERODIMER
CC WITH VEGF.
CC -1- SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR
CC TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN.
CC -1- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN HEART, BRAIN, KIDNEY
CC AND SKELETAL MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U48800; AAB06273.1; -
DR EMBL: U43837; AAC52553.1; -
DR HSSP: P15692; YVGH.
DR MGP: MGI:106199; Vegfb.
DR InterPro: IPR000072; -
DR Pfam: PF00341; PDGF_1;
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS00278; PDGF_2; 1.
KM Mitogen; Growth factor; Signal; Heparin-binding.
FT SIGNAL 1 21
FT CHAIN 22 188
FT SEQUENCE 188 AA; 21442 MW; D52A055FB9395E9CA CRC64;
SQ
Query Match 7.5%; Score 175; DB 1; Length 188;
Best Local Similarity 26.5%; Pred. No. 7.1e-07;
Matches 53; Conservative 27; Mismatches 80; Indels 40; Gaps 8;
OY 101 SRTETI-KFAAHNYTEIKSIDNEMRKTCQPREVCIDVGEKFGVATNTFFKPCVSV 159
DB 17 ARTQAAPVQFQDSRSHQKVVPWIDVYARAT-CQPREVVVPLSMELMGVNVQVPSCTIV 75
OY 160 YRCGCCNSGLQCMNTSTSYLSKTLFEITVPLSGPKPVYISFANHSCRCMSKLDVYR 219
DB 76 QRCGCCPPDDGLECVPTGHOVMQILMIRYPSQLGE---MSLEHSHQCECRPK----- 127
OY 220 QVHSIIRSLPATL-PQQAAN-----KTCPTNYMNNHICGLAODEPMFSSDAGDDST 273
DB 128 KESAVKPPDSPRLCPCTQRRRPPRTCR-----CQCRRRR----- 165

```

OY 274 DEFDICGPKLEDEETCC 293
 DB 166 --FLHCQGGRLNPDTCRC 183

RESULT 14

VEGF_SHEEP STANDARD: PRT: 146 AA.

AC P50412:
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR PRECURSOR (VEGF) (VASCULAR PERMEABILITY FACTOR) (VPF).
 GN VEGF.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Caprinae; Ovis.
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxId=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=97117958; PubMed=8958842;
 RA Redner D.A., Dai Y., Li J., Charnock-Jones D.S., Smith S.K., Reynolds L.P., Moor R.M.,
 RA "Characterization and expression of vascular endothelial growth factor (VEGF) in the ovine corpus luteum."
 RT J. Reprod. Fert. 108:157-165(1996).
 RL -

CC -1- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL CELL GROWTH. INDUCES ENDOTHELIAL PROLIFERATION AND VASCULAR PERMEABILITY.
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -1- SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X89506; CA61677.1; -
 DR HSSP: P15692; 1VPF.
 DR InterPro: IPR000072; -
 DR Pfam: PF00341; PDGF_1.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS00278; PDGF_2; 1.
 KW Mitogen; Growth factor; Glycoprotein; Signal.
 FT SIGNAL 1 26 BY SIMILARITY.
 FT CHAIN 27 146 VASCULAR ENDOTHELIAL GROWTH FACTOR.
 FT DISULFID 51 93 BY SIMILARITY.
 FT DISULFID 82 127 BY SIMILARITY.
 FT DISULFID 86 129 BY SIMILARITY.
 FT DISULFID 76 76 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 85 85 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 100 100 N-LINKED (GLCNAC...) (POTENTIAL).
 FT SEQUENCE 146 AA; 17247 MW; 4E792CB57F91760 CRC64;

Query Match 7.4%; Score 173.5; DB 1; Length 146;
 Best Local Similarity 29.0%; Pred.No. 7.2e-07;
 Matches 40; Conservative 18; Mismatches 61; Indels 19; Gaps 3;

OY 85 LKGGQHRREQANLSREELIKFAAHYNTLEIKSDENRKRQCPREVCIYGEKF 144
 DB 18 LHAHQAQAAAPAAEAGGQRPHEVTKPM-----DVGQRSCRPETLVDIFOEY 64

OY 145 GVATNTFFKPCSVYRCGCCNSSEGLQCMNTSTYLSKLTFLFIVPLSQ--GPKPVTS 202
 DB 65 PDIEFIKPCVPLMKRCGCCNDESLCEVPTPEFNITQIMRIKHOSHIG-----EMS 120
 OY 203 FANHTSCRCMSKLDVYRQ 220
 DB 121 FLQHNKCECRPKKDKARQ 138

RESULT 15

VEGF_PIG STANDARD: PRT: 2482 AA.

AC Q28833:
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE VON WILLEBRAND FACTOR PRECURSOR (VWF) (FRAGMENT).
 GN F8VWF OR VWF.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxId=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP Seaman W.T., Read M.S., Bellinger D.A., Nichols T.C.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 397-553 FROM N.A.
 RX MEDLINE=93356762; PubMed=8352759;
 RX Lavergne J.M., Piao Y.C., Ferreira V., Kerbiriou-Nabias D., Bahnak B.R., Meyer D.;
 RA "Primary structure of the factor VIII binding domain of human, porcine and rabbit von Willebrand factor."
 RT Biochem. Biophys. Res. Commun. 194:1019-1024(1993).
 RL -

CC -1- FUNCTION: IMPORTANT IN THE MAINTENANCE OF HOMEOSTASIS. IT PARTICIPATES IN PLATELET-VESSEL WALL INTERACTIONS BY FORMING A NONCOVALENT COMPLEX WITH COAGULATION FACTOR VIII AT THE SITE OF VASCULAR INJURY (BY SIMILARITY).
 CC -1- SUBUNIT: MULTIMERIC (BY SIMILARITY).
 CC -1- PTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 3 VWFc DOMAINS.
 CC -1- SIMILARITY: CONTAINS A C-TERMINAL CYSTEINE KNOT-LIKE DOMAIN (CTCK).
 CC -1- SIMILARITY: SOME, TO SILKMOB HEMOCYTIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AF052036; AAC06229.1; -
 DR EMBL: S64541; AAB27829.2; -
 DR HSSP: P04275; 1ATZ.
 DR InterPro: IPR000359; -
 DR InterPro: IPR001007; -
 DR InterPro: IPR001846; -
 DR InterPro: IPR002035; -
 DR InterPro: IPR002919; -
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00092; vwa; 3.
 DR Pfam: PF00093; vwc; 3.
 DR Pfam: PF00094; vwd; 3.
 DR Pfam: PF01826; tll; 3.
 DR PRINTS: PR00365; ENDOTHELIN.
 DR PRINTS: PR00453; VWFADOMAIN.
 DR PROSITE: PS00234; VWFc; 3.
 DR PROSITE: PS01208; VWFc; 3.
 DR PROSITE: PS01185; CTCK_1; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 KW Blood coagulation; Platelet; Glycoprotein; Extracellular matrix;

KW Plasma; Endothelial cell; Repeat; Cell adhesion.
FT NON_TER 1 1
FT PROPEP <1 437 BY SIMILARITY.
FT CHAIN 438 2482 VON WILLEBRAND FACTOR.
FT DOMAIN 62 215 WVFED 2.
FT DOMAIN 438 461 AMINO-TERMINAL.
FT DOMAIN 462 507 EL.
FT DOMAIN 500 527 CX.
FT DOMAIN 541 687 WVFED 3.
FT DOMAIN 947 1127 WVEA 1.
FT DOMAIN 1167 1334 WVEA 2.
FT DOMAIN 1360 1540 WVEA 3.
FT DOMAIN 1619 1771 WVEA 4.
FT DOMAIN 1885 1930 E2.
FT DOMAIN 1924 1997 WVEC 1.
FT DOMAIN 2098 2164 WVEC 2.
FT DOMAIN 2249 2319 WVEC 3.
FT DOMAIN 2393 2481 CRCK.
FT SITE 2176 2178 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 441 482 BY SIMILARITY.
FT DISULFID 450 478 BY SIMILARITY.
FT DISULFID 484 495 BY SIMILARITY.
FT DISULFID 541 670 BY SIMILARITY.
FT DISULFID 563 705 BY SIMILARITY.
FT DISULFID 572 667 BY SIMILARITY.
FT DISULFID 588 595 BY SIMILARITY.
FT DISULFID 734 758 BY SIMILARITY.
FT DISULFID 745 785 BY SIMILARITY.
FT DISULFID 763 765 BY SIMILARITY.
FT DISULFID 827 839 BY SIMILARITY.
FT DISULFID 823 843 BY SIMILARITY.
FT DISULFID 800 804 BY SIMILARITY.
FT DISULFID 870 873 BY SIMILARITY.
FT DISULFID 908 911 BY SIMILARITY.
FT DISULFID 942 1128 BY SIMILARITY.
FT DISULFID 1338 1339 BY SIMILARITY.
FT DISULFID 1358 1541 BY SIMILARITY.
FT DISULFID 1548 1573 BY SIMILARITY.
FT DISULFID 1568 1609 OR 1611 (BY SIMILARITY).
FT DISULFID 1641 1792 BY SIMILARITY.
FT DISULFID 1619 1754 BY SIMILARITY.
FT DISULFID 1596 1757 BY SIMILARITY.
FT DISULFID 1662 1670 BY SIMILARITY.
FT DISULFID 2393 2443 BY SIMILARITY.
FT DISULFID 2408 2457 BY SIMILARITY.
FT DISULFID 2419 2473 BY SIMILARITY.
FT DISULFID 2423 2475 BY SIMILARITY.
FT DISULFID ? 2480 BY SIMILARITY.
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 905 905 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1184 1184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1243 1243 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1640 1640 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1892 1892 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1959 1959 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2026 2026 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2069 2069 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2215 2215 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2254 2254 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2459 2459 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 2482 AA; 272394 MW; D499B7DDBFBCAEDD CRC64;

Query Match 7.0%; Score 163.5; DB 1; Length 2482;
Best Local Similarity 24.2%; Pred. No. 8e-05;
Matches 79; Conservative 25; Mismatches 113; Indels 109; Gaps 20;

OY 153 EKP-----CVSVYRCGGCNSGLQCMNTSTYSKTLFEITVPLSGPKPVITSF 203
DB 1905 FCPHQVLMGSCVPEEACTQCVDDGIR-----HGFLETVVPDQH---PCQI-- 1949
OY 204 ANHTGCRGMSKIDVYRQVHSITRSLPATLPGC-----QAANKTCPTVYMNHNHC 254

DB 1950 -----CTCLSG-----RRVNCITQPCTAPARAPAGLCIEVARLROEAHQCCP-----EYEC 1994
OY 255 RC-LAQEDF--MESSDAGDSTDGFDHICGPKNKEDEEFCOC---VCRAGLRPASCGRK 308
DB 1995 VCDLVSCDLPPYPHCEGGLQPLTLTNPGECPN-----FTCACRKECCPRGPLP-SCPPHR 2048
OY 309 E-----LDRNSCOCVCCKNKLFPSCGAGNREFDENTOCVCCKRTC----- 347
DB 2049 TPALRKTQCCDEYEACNCNVNTTISCPLGYLASTVINDCGCT-TTICLPDKVCVHRGTVY 2107
OY 348 PRNPPLNPGKACCECTESQKCLKGKRTKHOTCSYRPRCTNRQACERGFSY---SEE 404
DB 2108 PVGQFWEBCDVCCTHD-----LEDAYVGLRVACQACRRC---EDSCRPGFTYVLHEGE 2158
OY 405 VC-RCVP-----SYMK 414
DB 2159 CCGKCLPSACKVYIGSFRGDSYSYMK 2184

Search completed: October 17, 2001, 14:51:38
Job time: 464 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:43:54 ; Search time 128.06 Seconds
(without alignments)
432.889 Million cell updates/sec

Title: US-09-427-657-2

Perfect score: 2336

Sequence: 1 MHLLGFVSVCSLLAALP.....SYSEVRCVPSWKRPQMS 419

Scoring table: BLOSUM62

Gapop 10.0 ; Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP:archaea:*
2: SP:bacteria:*
3: SP:fungi:*
4: SP:human:*
5: SP:invertebrate:*
6: SP:mammal:*
7: SP:mhc:*
8: SP:organelle:*
9: SP:phage:*
10: SP:plant:*
11: SP:rodent:*
12: SP:unclassified:*
13: SP:vertebrate:*
14: SP:virus:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2092.5	89.6	420	6	Q9XS50 bos taurus
2	1819.5	77.9	418	13	O57352 coturnix co
3	704.5	30.2	354	4	O43915 homo sapien
4	703.5	30.1	358	11	P97946 mus musculu
5	675.5	28.9	326	11	O35251 rattus norv
6	634	27.1	126	11	O35757 rattus norv
7	269.5	11.5	1704	5	O94438 chironomus
8	252	10.8	1698	5	O94438 chironomus
9	238.5	10.2	232	4	O9H1W9 homo sapien
10	238.5	10.2	254	4	O16889 homo sapien
11	231	9.9	209	4	O60720 homo sapien
12	226.5	9.7	208	6	O9XS44 canis fami
13	226.5	9.7	214	6	O9MYV3 canis fami
14	221.5	9.5	214	6	O9XS44 canis fami
15	220.5	9.4	190	6	O9GKR0 equus caball
16	217.5	9.3	190	6	O9XS44 canis fami
17	216.5	9.3	190	6	O9GKR0 equus caball
18	212.5	9.1	190	11	O9QX39 spatax leuc
19	211	9.0	191	4	O75875 homo sapien

20	209	8.9	122	6	O9GLX1	O9GLX1 bos taurus
21	206.5	8.8	214	11	O90XG7	O90XG7 rattus norv
22	203.5	8.7	190	6	O7643	O7643 ovis aries
23	202	8.6	194	13	O42572	O42572 xenopus lae
24	200.5	8.6	174	4	O9UL23	O9UL23 homo sapien
25	195	8.3	171	4	O9H1W8	O9H1W8 homo sapien
26	194	8.3	704	3	O74567	O74567 trichoderma
27	192	8.2	142	11	O9ER6	O9ER6 mesocricetu
28	192	8.2	147	4	O9UH58	O9UH58 homo sapien
29	183.5	7.9	188	13	O73682	O73682 brachydantio
30	181	7.7	301	5	O9VWP6	O9VWP6 drosophila
31	179.5	7.7	148	13	O42571	O42571 xenopus lae
32	177	7.6	124	6	O9GK00	O9GK00 callithrix
33	174	7.4	999	4	O9W036	O9W036 homo sapien
34	173.5	7.4	118	6	O9WZB1	O9WZB1 ovis aries
35	173.5	7.4	150	11	O54881	O54881 rattus norv
36	173.5	7.4	188	6	O9XS48	O9XS48 bos taurus
37	173	7.4	800	5	O9VB78	O9VB78 drosophila
38	172.5	7.4	170	11	O9JXK7	O9JXK7 rattus norv
39	172.5	7.4	1810	13	O90824	O90824 gallus gall
40	170.5	7.3	146	11	O90XG6	O90XG6 rattus norv
41	165	7.1	144	13	O73822	O73822 brachydantio
42	164.5	7.0	765	5	O9ML50	O9ML50 sarcophaga
43	164	7.0	123	6	O9N1S1	O9N1S1 capreolus c
44	162	6.9	1511	5	O9VB21	O9VB21 drosophila
45	160.5	6.9	1532	13	O90994	O90994 gallus gall

ALIGNMENTS

RESULT 1
ID Q9XS50 PRELIMINARY: PRT: 420 AA.

AC Q9XS50: 01-NOV-1999 (TREMBL) 12, Created)
DT 01-NOV-1999 (TREMBL) 12, Last sequence update)
DT 01-MAR-2001 (TREMBL) 16, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEART;
RA Liu X., Yonekura H., Yamagishi S., Yamamoto Y., Yamamoto H.;
RT "Structure and expression of bovine VEGF family."
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB004275; BAA77687.1; -;
DR HSBP; P15692; IVP.
DR InterPro: IPR000072; -;
DR Pfam: PF00341; PDGF_1;
DR PROSITE: PS00249; PDGF_1; 1;
DR PROSITE: PS00278; PDGF_2; 1;
DR SMART: SM00141; PDGF_1;
KW signal.
FT SIGNAL. 1 20 POTENTIAL.
FT CHAIN 21 420 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
SQ SEQUENCE 420 AA; 46681 MW; 58BA84317A3C8E2D CRC64;

Query Match 89.6%; Score 2092.5; DB 6; Length 420;
Best local similarity 88.1%; Pred. No. 5.5e-19;
Matches 370; Conservative 19; Mismatches 30; Indels 1; Gaps 1;

QY 1 MHLLGFVSVCSLLAALPGREP-AAAAFESGDLSDAEPDAGERTAAASKLDEQ 59
DB 1 MHLLGFVSVCSLLAALPGREPAAAAFESGDLSDAEPDAGERTAAASKLDEQ 60
QY 60 LRSVSVDELMLVLPYEWKMKCOLRKGMQHNREQANLNSRTETTKFAAHNTETL 119
|||||

```

Db 61 LRSVSSDELMTVLYPEYWKMKCOLRKGMWSTEQANTINRTGTLFRAAHNYTEIL 120
Oy 120 KSIDNEMRKTOCMREVCIDVGEKFGVATNTFFKPCVSVYRCGCCNSBGLOCMNTSTY 179
Db 121 KSIDNEMRKTOCMREVCIDVGEKFGVATNTFFKPCVSVYRCGCCNSBGLOCMNTSTY 180
Oy 180 YLSKTLFEITVPLSOGPKPVITISFANHSTCRMSKLDVYRQVHSIIRSLPATLPQCGAA 239
Db 181 YLSKTLFEITVPLSOGPKPVITISFANHSTCRMSKLDVYRQVHSIIRSLPATLPQCGAA 240
Oy 240 NKTCPNVMNNHICRCLAQEDPFMSDAGDSTDFHICGPNKELDETCQVCRAGL 299
Db 241 NKTCPNVMNNHICRCLAQEDPFMSDAGDSTDFHICGPNKELDETCQVCRAGL 300
Oy 300 RPASCPHKLDRNSCQVCCKNKLFPSCGANCNEFDENTCQCCCKRTCPRNOLPNEGKA 359
Db 301 QASSCGPHKELDRNSCQVCCKNKLFPSSCGANCNEFDENTCQCCCKRTCPRNOLPNEGKA 360
Oy 360 CECTESPQKCLLKGGKFFHQTSCYRRPCTNRQKACEPGFSYSEEVYRCVPSYMKRPQMS 419
Db 361 CECTESPQKCLLKGGKFFHQTSCYRRPCTNRQKACEPGFSYSEEVYRCVPSYMKRPQMS 420

```

RESULT 2

```

ID 057352 PRELIMINARY: PRT: 418 AA.
AC 057352;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR.
VEGF-C.
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
NCBI_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98167900; PubMed=9435294;
RA Eichmann A., Corbel C., Jaffredo T., Breant V., Joukov V., Kumar V.,
Alitalo K., Le Douarin N.M.;
"Avian VEGF-C: cloning, embryonic expression pattern and stimulation
of the differentiation of VEGFR2-expressing endothelial cell
precursors.";
RT Development 125:743-752(1998).
RL EMBL: Y15837; CAA75799.1; -.
DR HSP: P15692; IYPP.
DR InterPro: IPR000072; -.
DR InterPro: IPR002400; -.
DR Pfam: PF00341; PDGF_1.
DR PRINTS: PR00438; GFCYSKNOT.
DR ProDom: PD001629; -. 1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS0278; PDGF_2; 1.
DR SMART: SM00141; PDGF_1.
DR Signal.
KM SIGNAL.
FT CHAIN 1 31 POTENTIAL.
FT CHAIN 111 418 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
SQ SEQUENCE 418 AA: 46839 MW: 099BFC79151BF2B CRC64;

```

Query Match 77.9%; Score 1819.5; DB 13; Length 418;
 Best Local Similarity 76.2%; Pred. No. 5.3e-165;
 Matches 320; Conservative 38; Mismatches 59; Indels 3; Gaps 3;

```

Oy 1 MHULGFFSVAGCLLAALPGPREAPAAAFESGLDSDAEPDAGEATVAASKDEQL 60
Db 1 MHULGFFSVAGCLLAALPGPREAPAAAFESGLDSDAEPDAGEATVAASKDEQL 59
Oy 61 RSVSSVDELMTVLYPEYWKMKCOLRKGMWSTEQANTINRTGTLFRAAHNYTEIL 120
Db 60 RSVSSVDELMTVLYPEYWKMKCOLRKGMWSTEQANTINRTGTLFRAAHNYTEIL 119

```

```

Oy 121 SIDNEMRKTOCMREVCIDVGEKFGVATNTFFKPCVSVYRCGCCNSBGLOCMNTSTY 180
Db 120 SIDNEMRKTOCMREVCIDVGEKFGVATNTFFKPCVSVYRCGCCNSBGLOCMNTSTY 179
Oy 181 LSKTLFEITVPLSOGPKPVITISFANHSTCRMSKLDVYRQVHSIIRSLPATLPQCGAA 240
Db 180 LSKTLFEITVPLSOGPKPVITISFANHSTCRMSKLDVYRQVHSIIRSLPATLPQCGAA 239
Oy 241 NKTCPNVMNNHICRCLAQEDPFMSDAGDSTDFHICGPNKELDETCQVCRAGL 299
Db 240 NKTCPNVMNNHICRCLAQEDPFMSDAGDSTDFHICGPNKELDETCQVCRAGL 298
Oy 300 RPASCPHKLDRNSCQVCCKNKLFPSCGANCNEFDENTCQCCCKRTCPRNOLPNEGKA 359
Db 299 RPASCPHKLDRNSCQVCCKNKLFPSSCGANCNEFDENTCQCCCKRTCPRNOLPNEGKA 358
Oy 360 CECTESPQKCLLKGGKFFHQTSCYRRPCTNRQKACEPGFSYSEEVYRCVPSYMKRPQMS 419
Db 359 CECTESPQKCLLKGGKFFHQTSCYRRPCTNRQKACEPGFSYSEEVYRCVPSYMKRPQMS 418

```

RESULT 3

```

ID 043915 PRELIMINARY: PRT: 354 AA.
AC 043915;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE GROWTH FACTOR FIGF.
GN FIGF OR VEGF-D.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98140120; PubMed=9479493;
RA Rodchigiani M., Lestingi M., Luddi A., Orlandini M., Franco B.,
Rossi E., Ballabio A., Zuffardi O., Oliviero S.;
"Human FIGF: cloning, gene structure, and mapping to chromosome Xp22.1
between the FIGA and the GFR genes.";
RT Genomics 47:207-216(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RX MEDLINE=97349118; PubMed=9205122;
RA Yamada Y., Nezu J., Shimane M., Hirata Y.;
"Molecular cloning of a novel vascular endothelial growth factor,
VEGF-D.";
RT Genomics 42:483-488(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98118549; PubMed=9435229;
RA Achen M.G., Jeltsch M., Kulk E., Maekinen T., Vitell A., Wilks A.F.,
Alitalo K., Stacker S.A.;
"Vascular endothelial growth factor D (VEGF-D) is a ligand for the
tyrosine kinases VEGF receptor 2 (Flk1) and VEGF receptor 3 (Flt4).";
RT Proc. Natl. Acad. Sci. U.S.A. 95:548-553(1998).
RL EMBL: Y12864; CAA73371.1; JOINED.
DR EMBL: Y12865; CAA73371.1; JOINED.
DR EMBL: Y12866; CAA73371.1; JOINED.
DR EMBL: Y12867; CAA73371.1; JOINED.
DR EMBL: Y12868; CAA73371.1; JOINED.
DR EMBL: Y12869; CAA73371.1; JOINED.
DR EMBL: Y12870; CAA73371.1; JOINED.
DR EMBL: D89630; BAA24264.1; -.
DR EMBL: AJ000185; CAA03942.1; -.
DR EMBL: Y12863; CAA73370.1; -.
DR HSP: P15692; IYPP.
DR InterPro: IPR000072; -.
DR Pfam: PF00341; PDGF_1.
DR ProDom: PD001629; -. 1.

```


RESULT	6		
035757			
ID	035757	PRELIMINARY;	PRT; 126 AA.
AC	035757;		
DT	01-JAN-1998 (TREMBLrel. 05, Created)		
DT	01-JAN-1998 (TREMBLrel. 05, last sequence update)		
DT	01-MAR-2001 (TREMBLrel. 15, last annotation update)		
DE	VASCULAR ENDOTHELIAL GROWTH FACTOR-C (FRAGMENT).		
OS	Rattus norvegicus (Rat).		
OC	Euarystota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Rattus.		
OX	NCBI_TaxId=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=SPRAGUE-DAWLEY; TISSUE=LUNG;		
RL	Mandriota S.J., Pepper M.S.;		
DR	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF010302; AAB63248.1; -.		
DR	HSSP; P15692; 2VPF.		
DR	InterPro; IPR000072; -.		
DR	Pfam; PF00341; PDGF.1.		
DR	PROSITE; PSS0278; PDGF_2.1.		
DR	SMART; SM00141; PDGF.1.		
FT	NON_TER	1	
FT	NON_TER	1	
FT	NON_TER	126	
SO	SEQUENCE	126 AA; 13977 MW; 8F365AFBC4E037B0 CRC64;	

RESULT	7	
Q94446		
ID	Q94446	PRELIMINARY;
AC	Q94446;	PRT; 1704 AA.
DT	01-FEB-1997	(TREMBlrel. 02, Created)
DT	01-FEB-1997	(TREMBlrel. 02, last sequence update)

Query Match	11.5%	Score 269.5	DB 5	Length 1704
Best Local Similarity	22.8%	Pred. No. 1e-16		
Matches 79	Conservative	48	Mismatches 123	Indels 97
				Gaps 16
QY 129	TCOMPREVCIDVGKEGVATNTFFKRPVCYSVVR-----CGGGCNSBGLCMNTSTSYLSK	183		
Db 1079	TNCPAKQOTITESECEGCET-----RKCLDGRFRSFLBEGCYCDEKKCY-----GK	1125		
QY 184	TFEFTIV-----PLS-----QGPRPYT-----ISFANHTSCRMS	213		
Db 1126	QIFDKNTKCKCKPNEKRGDSCGCKDFCVDCSKCKSPKPRANGCGVGEWHEMDCKCEC	1185		
QY 214	KLDVYQV-----HSIIRSLPATLQCOAAN-----KITPTNYM	248		
Db 1186	PKDKSTTGEGGQKMNDDNCCGCGCPPTATGASQKSYNVTSCGCGNPMGAPGPGNAV	1245		
QY 249	MNNHICRLCLAOEDFMFSSPADDDSTFDGFDHICGPNKLEIDETTCOCVCRAGLAPASGPHK	308		
Db 1246	WCENSCQCYCPKMEKPRADN-----CG-NKMMMDKACBECBCKGCEBAGCKGQY	1293		
QY 309	ELDRNSCCQVC-KNKLFPSCGANREEDENTQCYVCRT-----CPRNDPLNPGKCAECB	362		
Db 1294	KNNKNTCAECEPBGKATPASCGBDKKSMWDDSCQCKSKMPGGCGCPNQNMWCEFTCKKCC	1353		
QY 363	TESPOKCLLKGGKFFHHQTCSTCYRPPCTNOKKACEPFSISEVCRCV	409		
Db 1354	S-GTQTC-PAGOSWDSQTCG-SCPTGK---CTGAQFCAACQKCV	1394		

ID	094438	PRELIMINARY:	PRT:	1698 AA.
AC	094438:			
DT	01-FEB-1997	(TREMBLrel. 02, Created)		
DT	01-FEB-1997	(TREMBLrel. 02, Last sequence update)		
DT	01-OCT-2000	(TREMBLrel. 15, Last annotation update)		
DE	185 KDA SILK PROTEIN.			
GN	SP185.			
OS	Chironomus pallidivittatus (Midge).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;			
OC	Chironomidae; Chironomidae; Chironominae; Chironomus.			
OX	NCBI_TaxID=7151;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Salivary Gland;			
RA	Case S.T., Cox C., Bell W.C., Hoffman R.T., Martin J., Hamilton R.;			
RL	Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.			
DR	EMBL, U54640; AAA9803.1; -			
DR	InterPro: IPR0000561; -			
DR	InterPro: IPR000853; -			

PRINTS: PR00876; MTNEMATODE.
 DR PROSITE: PS00022; EGF 1; UNKNOWN 1.
 SO SEQUENCE 1698 AA; 186164 MW; 82AF0838D9E2360E CRC64;

Query Match 10.8%; Score 252; DB 5; Length 1698;
 Best Local Similarity 21.2%; Pred. No. 4.6e-15;

Matches 88; Conservative 40; Mismatches 123; Indels 164; Gaps 17;

123 DNEWR-----KTCMPREVCDIVGKEFGVATNTFFKPCVSYR-----CGGCCSEGL 171
 1161 DMQCKICPQVNTKCKDKQKFLSKCEGCDTQT-----QCKNGFRWSNLECCGLDEKRC 1116
 1172 QCMNNTSTYLSKLFETV-----PLS-----OGPKPVT-----I 201
 1117 Q-----GQVFDKNTCKCKCPKQKPDICGNGKDCPLDSCCKSPRANCGP 1167
 202 SFANHTSCRMCKLDVYRVHSIIRSLPATLPQCAANKTCPTNYMNNHICRLAOD 261
 1168 QEMNEKCCQ-----ECPKDKPK-----KCPGGQDMNNHLCQCCGCP 1206
 262 FHFSSAGDSDTDGFHDICGPNKELD-----ETCQCYCRAGL----- 299
 1207 APTCSNKQKYSNVSCSCGCKPKRPGKQIMCENTRCVCPKMDRPNANCGSKWMN 1266
 300 -----RPASGPHKELDRNSCOVCCKNKL 324
 1267 DMKQCECKPCEGCKCKVMMNANTCACCECPACKPACGAIKSNNDSCSCQCKSKMP 1326
 325 PSQCGANREDENTQCV-----KRTCPRNPLNPGKACCECTESPQ-----KCL 370
 1327 CGGCPNQQMNEKTCCECSAKGNCAPGOTWNSQTCQSCAPSGTCTGQVWCKACKCV 1386
 371 LKGR-----FHHOTGCYRPRCTNRK-----ACEGFSYSE-----EVCRCVP 410
 1387 CPAOKKCDSPKMTWDESSC-----QCPKNMRPPKGGCNGARTWDATCSERCAAVP 1438

RESULT 9
 Q9H1W9 PRELIMINARY: PRT; 232 AA.

AC Q9H1W9; 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE DJ261G23.6.3 (VASCULAR ENDOTHELIAL GROWTH FACTOR).
 GN VEGF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Williams S.;
 RL Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AL136131; CAC19512.1;
 SO SEQUENCE 232 AA; 27042 MW; F849F364446F4D01 CRC64;

Query Match 10.2%; Score 238.5; DB 4; Length 232;
 Best Local Similarity 24.7%; Pred. No. 8.6e-15;
 Matches 70; Conservative 33; Mismatches 101; Indels 79; Gaps 10;

69 LMTVLYPEYWKWKYKCOLRKGGQWQHNREQANLNSRTETIKFAAHYNTIELKSIDNEMRK 128
 14 LLLYLHNAKWSQAAPMAEGGQNNH-----EYVKEM-----DVGOR 49
 129 TQCMPEVDCIDGKEFGVATNTFFKPCVSYRGGCCNSELQCMNTSTYLSKTFEI 188
 50 SYCHPIETLVDFQEPDEIETIFKPCVPLMRGCGCDBEGLECVPEESITWQIMRI 109
 189 TVPLSOGKRPVTSIFANHTSCRMCKLDVYRVHSIIRSLPATLPQCAANKTCPTNYM 248

110 KP--HOGQHIGEMSFLOHNKCECPKADBARQEKKSVRG-----GKQKRRKRSRYKS 162
 249 WNNHT--CRCLAQEDPEFSSDADDDSDTDGFHDICGPNKELDEFTQCYCAGL--RPASGCP 306
 163 WSVYVGARC-----CLMPWSLPGPHPCGP 186
 307 HHE-----LDNSQCYCKNKLFPSCGANR--EFDENTCOC 341
 187 CSERKHLFVQDPQTCCKSCKNT--DSRCKARQELNERTCNC 227

RESULT 10
 Q16889 PRELIMINARY: PRT; 254 AA.

AC Q16889; 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 08, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR (FRAGMENT).
 GN VEGF 206.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92168017; PubMed=1791831;
 RA Houck K.A., Ferrara N., Winer J., Cachianes G., Li B., Leung D.W.;
 RT "The vascular endothelial growth factor family: identification of a
 RT fourth molecular species and characterization of alternative splicing
 RT of RNA."
 RL Mol. Endocrinol. 5:1806-1814(1991).
 DR EMBL: S85192; AAC63102.1;
 DR EMBL: S85224; AAC63101.1;
 DR EMBL: S85199; AAC63101.1; JOINED.
 DR EMBL: S85201; AAC63101.1; JOINED.
 DR EMBL: S85219; AAC63101.1; JOINED.
 DR EMBL: S85222; AAC63101.1; JOINED.
 DR HSPF: P15692; 2VPF.
 DR InterPro: IPR000072;
 DR Pfam: PF00341; PDGF_1.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS00278; PDGF_2; 1.
 DR SMART: SM00141; PDGF_1.
 FT NON-TER
 SO SEQUENCE 254 AA; 29461 MW; 069DFE9B97230BA8 CRC64;

Query Match 10.2%; Score 238.5; DB 4; Length 254;
 Best Local Similarity 24.7%; Pred. No. 9.6e-15;
 Matches 70; Conservative 33; Mismatches 101; Indels 79; Gaps 10;

69 LMTVLYPEYWKWKYKCOLRKGGQWQHNREQANLNSRTETIKFAAHYNTIELKSIDNEMRK 128
 36 LLLYLHNAKWSQAAPMAEGGQNNH-----EYVKEM-----DVGOR 71
 129 TQCMPEVDCIDGKEFGVATNTFFKPCVSYRGGCCNSELQCMNTSTYLSKTFEI 188
 72 SYCHPIETLVDFQEPDEIETIFKPCVPLMRGCGCDBEGLECVPEESITWQIMRI 131
 189 TVPLSOGKRPVTSIFANHTSCRMCKLDVYRVHSIIRSLPATLPQCAANKTCPTNYM 248
 132 KP--HOGQHIGEMSFLOHNKCECPKADBARQEKKSVRG-----GKQKRRKRSRYKS 184
 249 WNNHT--CRCLAQEDPEFSSDADDDSDTDGFHDICGPNKELDEFTQCYCAGL--RPASGCP 306
 185 WSVYVGARC-----CLMPWSLPGPHPCGP 208
 307 HKE-----LDNSQCYCKNKLFPSCGANR--EFDENTCOC 341
 209 CSERKHLFVQDPQTCCKSCKNT--DSRCKARQELNERTCNC 249

```

RESULT 11
ID 060720 PRELIMINARY: PRT: 209 AA.
AC 060720:
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE VEGF183 PROTEIN PRECURSOR (VASCULAR ENDOTHELIAL GROWTH FACTOR 183)
DE (D1261G23.6.6) (VASCULAR ENDOTHELIAL GROWTH FACTOR).
GN VEGF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=99096474; PubMed=9878851;
RA Lei J., Jiang A., Pei D.;
RT "Identification and characterization of a new splicing variant of
RT vascular endothelial growth factor: VEGF183."
RL Biochim. Biophys. Acta, Gene Struct. Expr. 1443:400-406(1998).
RN 12
RP SEQUENCE OF 114-209 FROM N.A.
RC TISSUE=RETINA;
RA Jijngling L., Roque R.S.;
RL Submitted (MAY-1998) to the EMBL/Genbank/DBJ databases.
RN 13
RP SEQUENCE FROM N.A.
RA Williams S.;
RL Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ010438; CA09179.1; -
DR EMBL: AF062645; AAC16730.1; -
DR EMBL: AL136131; CAC19514.1; -
DR HSSP: P15692; 2VPE.
DR InterPro: IPR000072; -
DR Pfam: PF00341; PDGF_1;
DR ProDom: PD001629; -; 1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS50278; PDGF_2; 1.
DR Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 209 VEGF183 PROTEIN.
SQ SEQUENCE 209 AA; 24422 MW; F01CCED945D6CA CRC64;

```

```

Query Match 9.7%; Score 231; DB 4; Length 209;
Best Local Similarity 24.2%; Pred. No. 4e-14; Indels 96; Gaps 9;
Matches 68; Conservative 28; Mismatches 87;

```

```

OY 69 LMTVLYPEYMKYKQOLRKGMQHNREOANLSRTEETIKFAAHYNTIELKSIDENMRK 128
DB 14 LLLYLHAKWSQAAPMAEGGQNNH-----EYVKRFM-----DYQGR 49
OY 129 TCCMREVCIIDVKEGVAATNTFFKPCVSVYRCGCCNSEGLQCMNTSTYLSKTLFEI 188
DB 50 SYCHPEITLVDIFOEYDEIEYIFKPCVPLMRGCCNDEGLCEVPTEESNITQIMRI 109
OY 189 TWPVLSGPRVITISFANHTSCRCMSKLDVYROYHSIIRSLPATLPQCAANKTCPTNYM 248
DB 110 KP--HOGGHIGKMSPLQHNKCECRKRDARQEKKSVR----- 145
OY 249 MNNHICRLAODEFMSSDAGDSTDFHDICGPNKELDEFTQCQVCAGLRPASCGRPK 308
DB 146 -----GKGGQKRR-----RKKSRR--CGPCS 165
OY 309 E-----LDNSCQCVCKNKLPPSQCAGNR-EFDENTCQC 341
DB 166 ERRKHLFVDDPQTKCKCKNT--DSRCKARQLELNERTCRC 204

```

RESULT 12
O9XSF4

```

ID O9XSF4 PRELIMINARY: PRT: 208 AA.
AC O9XSF4;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR 182.
GN VEGF.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedata; Canidae; Canis.
OX NCBI_TaxID=9615;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=HEART;
RA Jijngling L., Roque R.S.;
RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF133249; AAD29683.1; -
DR HSSP: P15692; 2VPE.
DR InterPro: IPR000072; -
DR Pfam: PF00341; PDGF_1;
DR ProDom: PD001629; -; 1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS50278; PDGF_2; 1.
DR SMART: SM00141; PDGF_1.
SQ SEQUENCE 208 AA; 24400 MW; CF77AC591F5C2BBE CRC64;

```

```

Query Match 9.7%; Score 226.5; DB 6; Length 208;
Best Local Similarity 24.4%; Pred. No. 1.1e-13;
Matches 66; Conservative 26; Mismatches 81; Indels 97; Gaps 9;

```

```

OY 85 LRKGQHNREOANLSRTEETIKFAAHYNTIELKSIDENMRKTCQMPREVCIDVKEF 144
DB 18 LHHAKWSQAAPMAEGGQNNH-----EYVKRFM-----DYQGRSVCRPIETLVDIFQY 64
OY 145 GVAATNTFFKPCVSVYRCGCCNSEGLQCMNTSTYLSKTLFEIYPL-----SQGPKV 199
DB 65 PDEIYIFKPCVPLMRGCCNDEGLCEVPTEE-----FNITQIMRIKPHOGGHIG 117
OY 200 TISFANHTSCRCMSKLDVYROYHSIIRSLPATLPQCAANKTCPTNYMNNHICRLAQ 259
DB 118 EMSFLQHSCECRKRDARQEKKSVR----- 144
OY 260 EDFMSSDAGDSTDFHDICGPNKELDEFTQCQVCAGLRPASCGRPK-----LDR 312
DB 145 -----GKGGQKRR-----RKKSRR--CGPCSERRKHLFVQDP 175
OY 313 NSCQCVCKNKLPPSQCAGNR-EFDENTCQC 341
DB 176 QTKCKCKNT--DSRCKARQLELNERTCRC 203

```

```

RESULT 13
ID O9MYV3 PRELIMINARY: PRT: 214 AA.
AC O9MYV3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR 188 PRECURSOR.
GN VEGF.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedata; Canidae; Canis.
OX NCBI_TaxID=9615;
RN 11
RP SEQUENCE FROM N.A.
RC MEDLINE=20125516; PubMed=10661874;
RA Scheidegger P., Weiglhofer W., Suarez S., Kaser-Holtz B., Steiner R.,
RA Ballmer-Hofer K., Jauszi R.;
RT "Vascular endothelial growth factor (VEGF) and its receptors in tumor-
RT bearing dogs."
RL Biol. Chem. 380:1449-1454(1999).

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:45:59 ; Search time 115.93 seconds
(without alignments)
53.339 Million cell updates/sec

Title: US-09-427-657-2_COPY_30_131
Perfect score: 539
Sequence: 1 AAFESGIDLSDAEPDAGEAT.....AHYNTILKSIDNEMRKTC 102

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues
Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
A.Geneseq_0601.*
1: /SIDSB/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SIDSB/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SIDSB/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SIDSB/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SIDSB/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SIDSB/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SIDSB/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SIDSB/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SIDSB/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SIDSB/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SIDSB/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SIDSB/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SIDSB/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SIDSB/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SIDSB/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SIDSB/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SIDSB/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SIDSB/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SIDSB/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SIDSB/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SIDSB/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SIDSB/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	539	100.0	399	20	AAW86237
2	539	100.0	419	18	AAW17837
3	539	100.0	419	18	AAW00932
4	539	100.0	419	18	AAW11478
5	539	100.0	419	19	AAW5751
6	539	100.0	419	19	AAW5740
7	539	100.0	419	20	AAW30518
8	539	100.0	419	20	AAW2320
9	539	100.0	419	20	AAW6203
10	539	100.0	419	21	AAW10648
11	539	100.0	419	21	AAW29048

12	539	100.0	419	21	AAW97144	Vascular endotheli
13	539	100.0	419	21	AAW70749	Human prepro-vascu
14	539	100.0	419	21	AAW70982	Human vascular end
15	539	100.0	419	22	AAW97570	Human VEGF-B prote
16	539	100.0	419	22	AAW37605	Human VEGF-C. Hom
17	530	98.3	419	18	AAW13833	Human vascular end
18	425	78.8	415	18	AAW00933	Mouse VEGF-2 prote
19	425	78.8	415	19	AAW5742	Mouse vascular end
20	412	76.4	418	18	AAW00934	Quail Flt4 recepto
21	412	76.4	418	19	AAW75743	Quail vascular end
22	347	64.4	350	20	AAW30519	A truncated vascular
23	347	64.4	350	20	AAW22321	Truncated human VE
24	347	64.4	350	21	AAW97145	Human VEGF-2 prote
25	347	64.4	350	22	AAW97577	Human VEGF-2 prote
26	342	63.5	350	16	AAW82686	Vascular endotheli
27	165	30.6	326	19	AAW44296	Rat vascular endot
28	163	30.2	178	20	AAW08287	Human growth facto
29	163	30.2	321	19	AAW53243	Mus musculus vascu
30	163	30.2	337	20	AAW08286	Human growth facto
31	163	30.2	358	19	AAW53242	Mus musculus vascu
32	163	30.2	358	19	AAW44295	Mouse vascular end
33	162	30.1	358	18	AAW14992	Marine c-Fos induc
34	161	29.9	113	20	AAW08285	Human growth facto
35	161	29.9	318	20	AAW08284	Human growth facto
36	155	28.8	325	19	AAW53240	Homo sapiens vascu
37	155	28.8	325	22	AAW97572	Human VEGF-D prote
38	155	28.8	354	19	AAW49036	Human vegf2 growt
39	155	28.8	354	19	AAW53241	Homo sapiens vascu
40	155	28.8	354	19	AAW4293	Human vascular end
41	155	28.8	354	21	AAW10649	Human VEGD protein
42	155	28.8	354	21	AAW29049	Human VEGF-D prote
43	155	28.8	354	21	AAW70750	Human prepro-vascu
44	155	28.8	354	21	AAW70983	Human vascular end
45	155	28.8	354	22	AAW97573	Human VEGF-D1 prot

ALIGNMENTS

RESULT 1	
AAW86237	standard; protein: 399 AA.
ID	AAW86237;
AC	AAW86237;
DE	16-FEB-1999 (first entry)
DT	Human VEGF-C full length sequence.
XX	
DE	VEGF: VRP; vascular endothelial growth factor; VEGF-related protein;
KW	recombinant; truncated; gene therapy; angiogenesis; cardiac ischaemia;
KW	coronary; collateral vessel development; cell growth; migration; heart;
KW	lower limb ischaemia; stroke; peripheral vascular disease; intestine;
KW	wound healing; skin; vascular permeability.
OS	Homo sapiens.
XX	
FN	W09849300-A2.
XX	
PD	05-NOV-1998.
XX	
PF	20-APR-1998; 98WO-US07801.
XX	
PR	25-APR-1997; 97US-0842984.
XX	
PA	(COLL-) COLLATERAL THERAPEUTICS.
XX	
PI	Bohlen P.
XX	
DR	WPI. 1999-009426/01.
XX	
PT	New truncated vascular endothelial growth factor-related protein
PT	subunits - lack part of the N-terminal sequence, used to stimulate

PT angiogenesis, e.g. for treating heart disease and ischaemia
XX
XX
PS Claim 5; Fig 2D; 113pp; English.
CC The invention relates to truncated VRP (vascular endothelial growth
CC factor (VEGF)-related protein) subunits that have at least one amino
CC acid N-terminal to the first Cys of the core sequence deleted. Host
CC cells transfected or transfected with expression vectors containing
CC nucleic acids encoding the truncated VRP subunits are used to produce
CC the truncated proteins recombinantly. The truncated VRP subunits,
CC optionally expressed from gene therapy vectors, have in vivo and in vitro
CC angiogenic activity and are used to stimulate angiogenesis, particularly
CC coronary collateral vessel development in cases of cardiac ischaemia; to
CC stimulate endothelial cell growth and migration in vitro; to treat heart
CC disease; to treat ischaemia (e.g. cardiac, chronic coronary or chronic
CC lower limb ischaemia), stroke and peripheral vascular disease); to promote
CC healing of wounds (of skin or intestines), and to increase vascular
CC permeability. Sequences AAW6234 to AAW6239 represent full length VRP
CC sequences from which the truncated fragments are created.
XX
XX
SQ Sequence 399 AA:
Query Match 100.0%; Score 539; DB 20; Length 399;
Best Local Similarity 100.0%; Pred. No. 1.1e-55;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AAFESGLDSDAPDAGEATAYASKDLEQRLSVSSVDELMVLYPEYKMKYCOLRKG 60
Db 10 aafesglidsdaepdaageatayaskdleeqrlsvssvdelmvtlypeykmkycqlrkqg 69
OY 61 WOHNRQOANINSTRTEETIKFAAAHYNTTELKSIDNEMRKTQC 102
Db 70 wqhnreganlnsrteetlkfaaahyntelklsidnewrtqgc 111
RESULT 2
ID AAW17837
AAW17837 standard; Protein; 419 AA.
XX
XX
AC AAW17837;
XX
DT 13-JAN-1998 (first entry)
XX
DE Human foetal liver kinase A binding protein flk-1bp.
XX
XX
KW Foetal liver kinase 1 binding protein; human; flk-1bp;
KW receptor tyrosine kinase; vasculogenesis; angiogenesis;
KW wound healing; tumour; therapy; antagonist; antibody.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT Peptide 1..20
FT /label= Sig-peptide
FT Protein 21..419
FT /label= Mat-protein
FT /note= "(Claim 10)"
FT Peptide 21..35
FT /label= N-terminal
FT /note= "(Claim 9)"
XX
XX
PN WO9717442-A1.
XX
XX
PD 15-MAY-1997.
XX
XX
PF 05-NOV-1996; 96WO-US17584.
XX
XX
PR 08-NOV-1995; 95US-0554374.
XX
XX
PA (IMMUNEX CORP.
XX
XX
PI Lyman SD;

XX
DR MPI: 1997-281031/25.
DR N-PSDB; AAT68811.
XX
XX
PT DNA encoding a human foetal liver kinase 1 binding protein - used
PT to treat conditions with insufficient protein, deliver agents to
PT cells and identify antagonists to treat protein-mediated conditions
XX
PS Claim 1; Page 30-32; 43pp; English.
XX
XX
CC This polypeptide comprises a human foetal liver kinase 1 binding
CC protein (flk-1bp) (see AAW17837) that binds to the receptor tyrosine
CC kinase flk-1 expressed on vascular endothelial and other cells.
CC The mature flk-1bp can be secreted from host cells transformed with
CC an expression vector including an isolated flk-1bp cDNA clone (see
CC AAT68811). Flk-1bp can be used to isolate cells to which it binds,
CC for use in studying the roles of such cells and of flk-1 in
CC vasculogenesis and angiogenesis. Angiogenesis inhibition or
CC increased vascularisation may be clinically desirable (e.g. to
CC suppress solid tumour growth or in wound healing, respectively).
CC The flk-1bp can be administered to treat conditions with defective
CC or insufficient flk-1. Polypeptides may also act as carriers to
CC deliver diagnostic/therapeutic agents to cells to which flk-1bp
CC binds, to generate antibodies, and to identify flk-1bp antagonists
CC useful for treating flk-1bp mediated conditions.
XX
XX
SQ Sequence 419 AA:
Query Match 100.0%; Score 539; DB 18; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.1e-55;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AAFESGLDSDAPDAGEATAYASKDLEQRLSVSSVDELMVLYPEYKMKYCOLRKG 60
Db 30 aafesglidsdaepdaageatayaskdleeqrlsvssvdelmvtlypeykmkycqlrkqg 89
OY 61 WOHNRQOANINSTRTEETIKFAAAHYNTTELKSIDNEMRKTQC 102
Db 90 wqhnreganlnsrteetlkfaaahyntelklsidnewrtqgc 131
RESULT 3
ID AAW00932
AAW00932 standard; Protein; 419 AA.
XX
XX
AC AAW00932;
XX
DT 10-NOV-1997 (first entry)
XX
XX
DE Human Flt4 receptor tyrosine kinase ligand VEGF-C.
XX
XX
KW VEGF-C; Flt4; receptor tyrosine kinase; VEGF-3; human;
KW vascular endothelial growth factor receptor-3; ligand;
KW angiogenesis; wound healing; lymph vessel; lymphangioma;
KW cancer; metastasis; therapy; diagnosis; antibody; inhibitor.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT Peptide 1..102
FT /label= Prepro-peptide
FT Peptide 32..227
FT /note= "preferred active fragment of VEGF-C,
FT retaining Flt4 ligand activity (Claim 15)"
FT Peptide 103..217
FT /note= "preferred active fragment of VEGF-C,
FT retaining Flt4 ligand activity (Claim 12)"
FT Peptide 103..225
FT /note= "preferred active fragment of VEGF-C,
FT retaining Flt4 ligand activity (Claim 13)"
FT Peptide 103..227
FT /note= "preferred active fragment of VEGF-C,

FT	Peptide	113..213	retaining Flt4 ligand activity (Claim 14)"
FT		/note="preferred active fragment of VEGF-C,	
FT		retaining Flt4 ligand activity (Claim 10)"	
FT	Peptide	113..227	
FT		/note="preferred active fragment of VEGF-C,	
FT		retaining Flt4 ligand activity (Claim 11)"	
FT	Peptide	131..211	
FT		/note="preferred active fragment of VEGF-C,	
FT		retaining Flt4 ligand activity (Claim 9)"	
FT	Peptide	161..221	
FT		/note="preferred active fragment of VEGF-C,	
FT		retaining Flt4 ligand activity (Claim 8)"	
PN	WO9705250-A2.		
XX	13-FEB-1997.		
PD			
XX	01-AUG-1996;	96WO-FI00427.	
XX			
XX	28-JUN-1996;	96US-0671573.	
PR	01-AUG-1995;	95US-0510133.	
PR	12-JAN-1996;	96US-0585895.	
PR	14-FEB-1996;	96US-0601132.	
XX	(UYHE-) UNIV HELSINKI LICENSING LTD OY.		
XX			
PI	Alltalo K, Joukov V;		
XX			
DR	WPI: 1997-145688/13.		
XX	N-PSDB; AAT84276.		
PT	Flt4 receptor tyrosine kinase ligand and related nucleic acid - used		
PT	to modulate growth of endothelial cells and for diagnosis of		
PT	endothelial cell diseases		
PS	Claim 7; Page 112-113; 183pp; English.		
XX			
CC	This polypeptide comprises the pre-pro sequence of human VEGF-C,		
CC	a novel ligand that binds specifically to human Flt4 receptor.		
CC	Tyrosine kinase (VEGFR-3), stimulating phosphorylation of the		
CC	receptor. Its sequence was deduced from a cDNA clone (AAT84276)		
CC	obt. from a PC-3 prostatic adenocarcinoma cell (ATCC CRL 1435)		
CC	library. The polypeptide, or its active fragments, can be		
CC	expressed in transformed or transfected host cells for use in		
CC	claimed methods for detecting endothelial cells (e.g. to image		
CC	lymphatic vessels, endothelial venules), Flt4 receptor in		
CC	histochemical tissue) and also to modulate the growth of mammalian		
CC	endothelial cells (e.g. to accelerate angiogenesis and to promote		
CC	endothelial function of lymphatic vessels). Inhibitors of		
CC	VEGF-C, such as antibodies, can be used to control endothelial		
CC	cell proliferation, e.g. lymphangioma or metastatic cancer.		
CC	Mouse and quail VEGF-C sequences (see AAM00934-35) have also been		
CC	isolated.		
XX			
XX	Sequence 419 AA;		
XX			

Query Match	100.0%;	Score 539;	DB 18;	Length 419;
Best Local Similarity	100.0%;	Pred. No. 1,1e-55;		
Matches 102;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	AAFEGLDLSDAEPDAGEATAYASKDLEEQLRSSVSDLMVTLYPEYKMYKCOLRKGG	60	
Db	30	aafegldlsddepagatataskdlseqrtrsvsdvdelmtvlypeykwmykcoqlrkcg	89	
QY	61	WQHNREGANLNSRTDEETIKFAAHNTETELKSIDNEMKRTQC	102	
Db	90	wqnhreaganlnsrteetlikfaahnyntetelksidnewrtqcg	131	
RESULT	4			
AAAIL478				

ID	AAW11478 standard; Protein; 419 AA.
XX	
AC	AAW11478;
XX	
DT	23-APR-1997 (first entry)
DE	
XX	Human vascular endothelial growth factor 2.
XX	
KW	Vascular endothelial growth factor 2; VEGF2; angiogenesis;
KW	endothelialisation; coronary bypass surgery; vascular graft surgery;
KW	agonist; antagonist; therapy; diagnosis.
XX	
OS	Homo sapiens.
XX	
EH	Key
FT	Peptide
FT	1..46
FT	/label= Sig_peptide
FT	47..419
FT	/label= Mat.protein
FT	/note= "the mature protein is separately claimed
FT	(Claim 5)"
XX	
PN	W09639515-A1.
XX	
PD	12-DEC-1996.
XX	
PE	06-JUN-1996; 96MO-US09001.
XX	
PR	06-JUN-1995; 95US-0465968.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Cao L, Hu J, Rosen CA;
XX	
DR	WPI; 1997-043137/04.
XX	
XX	N-PSDB; AAT51371.
PT	
PS	Claim 1; Fig 2; 74pp; English.
XX	
CC	Human vascular endothelial growth factor 2 (VEGF2) (AAW11478) is
CC	structurally related to the VEGF/PDGF family and is a potent
CC	mitogen for vascular endothelial cells, stimulating their growth
CC	and angiogenesis. The amino acid sequence of VEGF2 was deduced
CC	from a cDNA clone (AAT51371) obtd. from an early stage human (week 9
CC	embryo) cDNA library. VEGF2 polypeptides can be produced in
CC	transformed host cells and used to promote angiogenesis e.g., to
CC	stimulate the growth of transplanted tissue following coronary
CC	bypass surgery, or to promote endothelialisation in vascular graft
CC	surgery. It can also be used to screen for antagonists (useful
CC	e.g. for tumour therapy) and agonists of VEGF2 activity.
XX	
Sequence	419 AA.
50	

Query Match	100.0%	Score 539;	DB 18;	Length 419;
Best Local Similarity	100.0%	Pred. No. 1,1e-55;		
Matches 102;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	AAFEGLDLSDAEPDAGEATAYASKDLSEQLSVSSVDLMTLVLPYWKMYKCOLRKGG	60	
Db	30	aafegldlsdepdageatayaskdlseqlrsvssvdelmtlvlpwykmykqqlrkgy	89	
QY	61	WQHNREQANLSRTEETIKFPAAHYNTETILKSIDENMRKTCQ	102	
Db	90	wqhnreqanlsrteetlkfaahyntetilsidenwrktgc	131	
RESULT	5			
ID	AAW75751	standard; Protein: 419 AA.		

```
XX AAW75751;
AC 14-DEC-1998 (first entry)
DT
DE Vascular endothelial growth factor C protein analogue.
XX
XX F1t4; Vascular endothelial growth factor C; vascular endothelial cell;
XX Lymphatic endothelial cell; myelopoiesis; angiogenesis; inflammation;
KM lymphangiogenesis; oedema; elephantiasis; Milroy's disease.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Modified-site 156
FT /note= "Xaa can be anything other than cysteine, or
FT can be nothing"
XX
XX W09833917-A1.
XX
XX 06-AUG-1998.
XX
XX 02-FEB-1998; 98MO-US01973.
XX
XX 05-FEB-1997; 97US-0795430.
XX
XX (LUDM-) LUDMIG INST CANCER RES.
XX (UYHE-) UNIV HELSINKI LICENSING LTD.
XX
XX Alitalo K, Joukov V;
XX
XX WPI; 1998-437470/37.
XX
XX New isolated vascular endothelial growth factor polypeptide(s) -
PT used to develop products for treating, e.g. cancers, inflammation,
PT oedema, granulocytopenia or for wound healing or tissue
PT transplantation
XX
XX Example 35; Page 143-145; 177pp; English.
XX
XX The vascular endothelial growth factor C (VEGF-C) polypeptides have
XX activities affecting growth and migration of vascular endothelial cells,
XX promoting growth of lymphatic endothelial cells and lymphatic vessels,
XX increasing vascular permeability, and affecting myelopoiesis. The
XX products can be used for stimulating angiogenesis, for inhibiting
XX angiogenesis, for stimulating lymphangiogenesis, treatment or prevention
XX of inflammation, oedema, elephantiasis, or Milroy's disease. They can
XX also be used to modulate myelopoiesis, e.g. treating granulocytopenia.
XX They can also be used for modulating the growth of endothelial cells.
XX They can also be used to stimulate lymphocyte production and maturation,
XX and to promote or inhibit trafficking of leucocytes between tissues and
XX lymphatic vessels or to affect migration in and out of the thymus.
XX
XX Sequence 419 AA:
```

```
Query Match 100.0%; Score 539; DB 19; Length 419;
Best Local Similarity 100.0%; Pred. No. 1,1e-55;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AAFESGLDSDAEPDAGATAYASKDLEQRLSVSSVDLMVLYPEYKMYKCOLRRKG 60
DB 30 aafesglidsdepdagetayaskdleqrlsvssvdelmvllypeykmkcolrrkg 89
QY 61 WQHNREQANLNRTEETIKFAAAHYNTETILKSINDENRRTQC 102
DB 90 wqhnreqanlnsrteetlkfaaahyntetllksidnewrktgc 131
```

```
RESULT 6
AAW75740
ID AAW75740 standard; Protein; 419 AA.
XX
```

```
AC AAW75740;
XX
XX 20-NOV-1998 (first entry)
DT
DE Human vascular endothelial growth factor C protein.
XX
XX F1t4; Vascular endothelial growth factor C; vascular endothelial cell;
XX Lymphatic endothelial cell; myelopoiesis; angiogenesis; inflammation;
KM lymphangiogenesis; oedema; elephantiasis; Milroy's disease.
XX
XX Homo sapiens.
OS
XX
XX W09833917-A1.
XX
XX 06-AUG-1998.
XX
XX 02-FEB-1998; 98MO-US01973.
XX
XX 05-FEB-1997; 97US-0795430.
XX
XX (LUDM-) LUDMIG INST CANCER RES.
XX (UYHE-) UNIV HELSINKI LICENSING LTD.
XX
XX Alitalo K, Joukov V;
XX
XX WPI; 1998-437470/37.
XX
XX N-PSDB; AAW52576.
XX
XX New isolated vascular endothelial growth factor polypeptide(s) -
PT used to develop products for treating, e.g. cancers, inflammation,
PT oedema, granulocytopenia or for wound healing or tissue
PT transplantation
XX
XX Claim 1; Page 112-115; 177pp; English.
XX
XX The vascular endothelial growth factor C (VEGF-C) polypeptides have
XX activities affecting growth and migration of vascular endothelial cells,
XX promoting growth of lymphatic endothelial cells and lymphatic vessels,
XX increasing vascular permeability, and affecting myelopoiesis. The
XX products can be used for stimulating angiogenesis, for inhibiting
XX angiogenesis, for stimulating lymphangiogenesis, treatment or prevention
XX of inflammation, oedema, elephantiasis, or Milroy's disease. They can
XX also be used to modulate myelopoiesis, e.g. treating granulocytopenia.
XX They can also be used for modulating the growth of endothelial cells.
XX They can also be used to stimulate lymphocyte production and maturation,
XX and to promote or inhibit trafficking of leucocytes between tissues and
XX lymphatic vessels or to affect migration in and out of the thymus.
XX
XX Sequence 419 AA:
```

```
Query Match 100.0%; Score 539; DB 19; Length 419;
Best Local Similarity 100.0%; Pred. No. 1,1e-55;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AAFESGLDSDAEPDAGATAYASKDLEQRLSVSSVDLMVLYPEYKMYKCOLRRKG 60
DB 30 aafesglidsdepdagetayaskdleqrlsvssvdelmvllypeykmkcolrrkg 89
QY 61 WQHNREQANLNRTEETIKFAAAHYNTETILKSINDENRRTQC 102
DB 90 wqhnreqanlnsrteetlkfaaahyntetllksidnewrktgc 131
```

```
RESULT 7
AAW30518
ID AAW30518 standard; Protein; 419 AA.
XX
XX AAW30518;
XX
XX 16-NOV-1999 (first entry)
DT
XX Vascular endothelial growth factor-2 (VEGF-2).
XX
```

Human vascular endothelial growth factor-2; VEGF-2; vascular endothelial cell growth; endothelial cell migration; angiogenesis; blood pressure; blood flow; immune system disorder; immune cell; cancer; autoimmune disorder; blood protein disorder; ataxia telangiectasia; common variable immunodeficiency; DiGeorge syndrome; HIV infection; HTLV-BLV infection; leukocyte adhesion deficiency syndrome; lymphopenia; phagocyte bactericidal dysfunction; severe combined immunodeficiency; Miskott-Aldrich disorder; anemia; thrombocytopenia; hemoglobinuria; allergy; asthma; allergic asthma.

Homo sapiens.

WO9946364-A1.

16-SEP-1999.

10-MAR-1999; 99WO-US05021.

13-MAR-1998; 98US-0042105.

30-JUN-1998; 98US-0107997.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Cao L, Hu J;

WPI; 1999-551399/46.

N-PSDB; AA210523.

New human vascular endothelial growth factor-2, used for treating, e.g. immune disorders and cancers

Claim 12; Fig 1A-E; 222pp; English.

The present sequence represents vascular endothelial growth factor-2 (VEGF-2). The VEGF-2 polypeptides have activities similar to VEGF. The VEGF-2 polypeptides stimulate the growth of vascular endothelial cells, stimulate endothelial cell migration, stimulate angiogenesis, decrease blood pressure, and increase blood flow. The polynucleotides and polypeptides can be used for preventing, treating or ameliorating a medical condition. The VEGF-2 polypeptides or polynucleotides may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or toxins), or infectious. Examples of immunologic deficiency syndromes include blood protein disorders, ataxia telangiectasia, common variable immunodeficiency, DiGeorge syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDS), Miskott-Aldrich disorder, anemia, thrombocytopenia, or hemoglobinuria. They can also be used to modulate emostatic or thrombolytic activity. Similarly allergic reactions and conditions such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated.

```

Query Match Similarity      100.0%: Score 539: DB 20: Length 419:
Best Local Similarity      100.0%: Pred. NO. 1,1e-55:
Matches 102: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY      1  AAFESGLDLSDAEPDAGATAYASNDLEQLSVSSVDLMTLVLPKYWKMKCOLRKGG 60
      |||||
Db       30  aafesgltdsdepdagatataskdleeqlrsvssvdeimtlvlpkywkmykqlrkgg 89
      |||||

QY      61  WQHNREQANLNSRTEETIKFAAAHYNTETILKSIDNEMRKTQC 102
      |||||
Db       90  wqhnreaganlnsrteetikfaahynteilksidnewrktgc 131
      |||||

```

RESULT 8
 AAY22320
 ID AAY22320 standard; Protein; 419 AA.
 XX
 AC AAY22320;
 XX
 DT 22-SEP-1999. (first entry)
 XX
 DE Full length human VEGF2 protein sequence.
 XX
 KW VEGF2; vascular endothelial growth factor 2; angiogenesis; bone damage;
 RW endothelial cell proliferation; tissue damage; therapy.
 XX
 OS Homo sapiens.
 XX
 PN US5932540-A.
 XX
 PD 03-AUG-1999.
 XX
 PE 24-DEC-1997; 97US-0999811.
 XX
 PR 24-DEC-1997; 97US-0999811.
 PR 08-MAR-1994; 94US-0207550.
 PR 06-JUN-1995; 95US-0465968.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Gao L, Hu J, Rosen CA.
 XX
 DR WPI; 1999-443606/37.
 DR N-PSDB; AAX84837.
 XX
 PT Vascular endothelial growth factor 2 for wound healing and vascular
 PT repair
 XX
 CL Claim 1; Fig 1; 49pp; English.
 CC
 CC This sequence is the vascular endothelial growth factor 2 (VEGF2),
 CC of the invention. The isolated polypeptide is useful for stimulating
 CC angiogenesis, by promoting the proliferation of endothelial cells, for
 CC the treatment of a wound, or for the treatment of tissue or bone damage.
 CC
 SQ Sequence 419 AA;

	Query Match	100.0%; Score 539; DB 20; Length 419;
	Best Local Similarity	100.0%; Pred. NO. 1.le-55;
	Matches 102; Conservative	0; Mismatches 0; Indels 0; Gaps 0
QY	1 AAFESGLDLSDEPPAGATAYASNDLEFOLSVSSVDLMTLVLYPEYKMYKCQLRRKG 60 	
Dd	30 aafesglstdeppagatayaskdlleqlrsvssvdselmtlvlypeykmkqqlrrkg 89 	
QY	61 WQHNRQANLSRTEETIKFAAHYNTELKSIDNEMRRKTQC 102 	
Dd	90 wqhnrqanlnsrteetlkfaahynntellksidnewrkrtgc 131 	
RESULT	9	
AAW86203		
ID	AAW86203 standard; protein; 419 AA.	
XX		
AC	AAW86203;	
XX		
DT	16-FEB-1999 (first entry)	
XX		
DE	Human vascular endothelial growth factor (VEGF)-C sequence.	
XX		
KM	VEGF; VRP; vascular endothelial growth factor; VEGF-related protein; recombinant; truncated; gene therapy; angiogenesis; cardiac ischemia; coronary; collateral vessel development; cell growth; migration; heart; lower limb ischemia; stroke; peripheral vascular disease; intestine; wound healing; skin; vascular permeability.	

```

XX OS Homo sapiens.
XX PN MO9849300-A2.
XX PD 05-NOV-1998.
XX PF 20-APR-1998; 98WO-US07801.
XX PR 25-APR-1997; 97US-0842984.
XX PA (COLL-) COLLATERAL THERAPEUTICS.
XX PI Bohlen P.
XX DR WPI: 1999-009426/01.
XX PT New truncated vascular endothelial growth factor-related protein
XX subunits - lack part of the N-terminal sequence, used to stimulate
XX angiogenesis, e.g. for treating heart disease and ischaemia
XX PS
XX PS Disclosure: Fig 1; 113pp: English.
XX CC This represents the amino acid sequence of human vascular endothelial
XX growth factor (VEGF)-C protein. The invention provides truncated VRF
XX (VEGF-related protein) subunits that have at least one amino acid
XX C-terminal to the first Cys of the core sequence deleted. Host cells
XX transformed or transfected with expression vectors containing nucleic
XX acids encoding the truncated VRF subunits are used to produce the
XX truncated proteins recombinantly. The truncated VRF subunits, optionally
XX expressed from gene therapy vectors, have in vivo and in vitro angiogenic
XX activity and are used to stimulate angiogenesis, particularly coronary
XX collateral vessel development in cases of cardiac ischaemia: to stimulate
XX endothelial cell growth and migration in vitro; to treat heart disease;
XX to treat ischaemia (e.g. cardiac, chronic coronary or chronic lower limb
XX ischaemia; stroke and peripheral vascular disease); to promote healing of
XX wounds (of skin or intestines), and to increase vascular permeability.
XX SQ Sequence 419 AA:

Query Match 100.0%; Score 539; DB 20; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.1e-55;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAFESGLDISDAEPDAGEATAYASKDLEQRLRSVSSVDELMVTLYPEYKMKCOLRKGG 60
DB 30 aafesgldisdapdagatayaskdleeqlrsvssvdelmtvlypeykmkqqlrksg 89
OY 61 WOHNRQANLNSRTEETIKFAAAHYNTETIKSIDNEMRKTOC 102
DB 90 wqhnrqanlnsrteetlkfaaahyntetlksidnewrtkgtc 131

RESULT 10
AAB10648
ID AAB10648 standard; Protein; 419 AA.
XX AC AAB10648;
XX DT 19-JAN-2001 (first entry)
XX DE Human VEGC protein.
XX DE VEGF-X; vascular endothelial growth factor; human; vulnery; cytosstatic;
XX anti-rheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
XX angiogenesis regulator; vascularization regulator; cancer; psoriasis;
XX rheumatoid arthritis; vascular retinopathy; blood vessel; organ repair;
XX tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
XX venous sore; diabetic ulcer; burns; skin graft growth; VEGC.
XX OS Homo sapiens.
XX XX

```

```

PN WO200037641-A2.
XX PD 29-JUN-2000.
XX PF 21-DEC-1999; 99WO-US30503.
XX PR 22-DEC-1998; 98GB-0028377.
XX PR 18-MAR-1999; 99US-0124967.
XX PR 08-NOV-1999; 99US-0164131.
XX PA (JANC ) JANSSEN PHARM NV.
XX PI Gordon RD, Sprengel JJ, Yon JR, Dijkmans JTH, Gosielska A;
XX PI Dhanaraj SN, Xu J;
XX DR WPI: 2000-442669/38.
XX PT New vascular endothelial growth factor protein, useful for treating or
XX preventing diseases associated with inappropriate angiogenesis activity
XX such as cancer, rheumatoid arthritis, psoriasis and wounds -
XX PS
XX PS Disclosure: Fig 11; 127pp; English.
XX CC This invention describes a novel vascular endothelial growth factor-X
XX (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
XX vulnery, cytosstatic, anti-rheumatic, antiarthritic, antipsoriatic and
XX antidiabetic activity and acts as an angiogenesis and vascularization
XX regulator. An antisense molecule of the invention is useful for treating
XX or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
XX retinopathy by inhibiting angiogenic activity or inappropriate
XX vascularization including formation and proliferation of new blood
XX vessels, growth and development of tissues, tissue regeneration and organ
XX and tissue repair in a subject. The products of the invention are useful
XX for preparing medicaments for treating wounds such as dermal ulcers,
XX pressure sores, venous sores, diabetic ulcers and burns and to promote
XX skin graft growth, tissue repair, proliferation of new blood vessels,
XX tissue regeneration and organ repair by promoting angiogenic activity or
XX vascularization. This sequence represents the human VEGC protein used
XX to illustrate the method of the invention.
XX SQ Sequence 419 AA:

Query Match 100.0%; Score 539; DB 21; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.1e-55;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAFESGLDISDAEPDAGEATAYASKDLEQRLRSVSSVDELMVTLYPEYKMKCOLRKGG 60
DB 30 aafesgldisdapdagatayaskdleeqlrsvssvdelmtvlypeykmkqqlrksg 89
OY 61 WOHNRQANLNSRTEETIKFAAAHYNTETIKSIDNEMRKTOC 102
DB 90 wqhnrqanlnsrteetlkfaaahyntetlksidnewrtkgtc 131

RESULT 11
AAB29048
ID AAB29048 standard; Protein; 419 AA.
XX AC AAB29048;
XX DT 31-JAN-2001 (first entry)
XX DE Human VEGF-C protein sequence.
XX DE Human; Flt4; fms-like tyrosine kinase 4; lymphoedema;
XX vascular endothelial growth factor receptor 3; VEGFR-3;
XX Milroy-Nonne syndrome; lymphoedema praecox; VEGF-C;
XX vascular endothelial growth factor C.
XX OS Homo sapiens.
XX XX

```


PN WO200058511-A1.
 XX
 PD 05-OCT-2000.
 XX
 XX 26-MAR-1999; 99WO-US06133.
 PF 26-MAR-1999; 99WO-US06133.
 XX
 XX 26-MAR-1999; 99WO-US06133.
 PR
 XX (LUDWIG-) LUDWIG INST. CANCER RES.
 PA (UYRE-) UNIV. HELSINKI LICENSING LTD OY.
 PA (UYPI-) UNIV. PITTSBURGH.
 XX
 PI Ferrell RE, Allitalo K, Finegold DN, Karkkainen M;
 XX
 XX WPI: 2000-679298/66.
 DR N-PSDB: AAC62406.
 XX
 XX Screening a human subject for increased risk of developing a lymphatic
 PT disorder, comprises assaying a nucleic acid to determine a mutation
 PT altering the sequence of a vascular endothelial growth factor
 PT receptor-3 -
 XX
 PS Disclosure: Page 60-61; 76pp; English.
 XX
 CC The present sequence is the protein sequence for the human vascular
 CC endothelial growth factor C (VEGF-C). It was used to demonstrate the
 CC methods of the invention, which involve the screening of individuals to
 CC determine which vascular endothelial growth factor receptor 3 (VEGFR-3,
 CC also known as Flt4 or fms-like tyrosine kinase 4) alleles they possess
 CC and thus their likelihood of developing hereditary lymphoedema.
 CC Conditions associated with lymphoedema include Milroy-Nonne syndrome,
 CC which is early onset lymphoedema and lymphoedema praecox, which is late
 CC onset.
 CC
 SQ Sequence 419 AA:
 Query Match 100.0%; Score 539; DB 21; Length 419;
 Best Local Similarity 100.0%; Pred. No. 1,1e-55;
 Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAFESGLDISDAEPDAGETATYASKDLEQLRSVSSVDELMTVLYPEYKMKCOLRKGG 60
 DB 30 aafesglidisdaepdagetaayaskdleqlrsvasvdelmtvlypeykmkcolrkqlyg 89
 QY 61 WQHNREQANLNSRTEETIKFAAHYNTETIKSIDNEMKRTQC 102
 DB 90 wqhnreganlnsrteetlkfaaahyntelkksidnewrtqtc 131
 RESULT 12
 AAY97144
 ID AAY97144 standard; Protein: 419 AA.
 XX
 XX AAY97144;
 AC
 XX 22-DEC-2000 (first entry)
 DT
 XX
 DE Vascular endothelial growth factor-2 (VEGF-2).
 XX
 KW Vascular endothelial growth factor 2; VEGF-2; retina; angiogenesis;
 KW treatment; injury; degeneration; photoreceptors; eye;
 KW angiod streaks; retinitis; pigmentosa; human;
 KW age-related macular degeneration; diabetic retinopathy.
 XX
 OS Homo sapiens.
 XX
 PN WO200045835-A1.
 XX
 PD 10-AUG-2000.
 XX
 PF 07-FEB-2000; 2000WO-US03047.
 XX

PR 08-FEB-1999; 99US-0119179.
 PR 12-FEB-1999; 99US-0119926.
 PR 03-JUN-1999; 99US-0137796.
 PR 22-DEC-1999; 99US-0171505.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Alderson R, Meider R, Roschke V, Ruben SM;
 XX
 XX WPI: 2000-532862/48.
 DR N-PSDB: AAA52080.
 XX
 XX Treating injury or degeneration of photoreceptors comprises
 PT administering to a subject vascular endothelial growth factor 2
 PT (VEGF-2)
 XX
 PS Claim 31; Fig 1a-e; 252pp; English.
 XX
 CC Administration of vascular endothelial growth factor 2 (VEGF-2)
 CC to a patient can be used for treating injury or degeneration of
 CC photoreceptors associated with e.g. angiod streaks, retinitis
 CC pigmentosa, age-related macular degeneration, diabetic retinopathy,
 CC etc. VEGF-2 promotes angiogenesis, the formation of new blood
 CC vessels in the retina.
 CC
 SQ Sequence 419 AA:
 Query Match 100.0%; Score 539; DB 21; Length 419;
 Best Local Similarity 100.0%; Pred. No. 1,1e-55;
 Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAFESGLDISDAEPDAGETATYASKDLEQLRSVSSVDELMTVLYPEYKMKCOLRKGG 60
 DB 30 aafesglidisdaepdagetaayaskdleqlrsvasvdelmtvlypeykmkcolrkqlyg 89
 QY 61 WQHNREQANLNSRTEETIKFAAHYNTETIKSIDNEMKRTQC 102
 DB 90 wqhnreganlnsrteetlkfaaahyntelkksidnewrtqtc 131
 RESULT 13
 AAY70749
 ID AAY70749 standard; Protein: 419 AA.
 XX
 XX AAY70749;
 AC
 XX 17-AUG-2000 (first entry)
 DT
 XX
 DE Human prepro-vascular endothelial growth factor C.
 XX
 KW Human; receptor tyrosine kinase; RTK; Flt4; fms-like tyrosine kinase 4;
 KW VEGFR-3; vascular endothelial growth factor receptor-3; chromosome 5q35;
 KW cyrostatic; tumour imaging; anti-tumour therapy; treatment; diagnosis;
 KW neoplastic disease; lymphoma; carcinoma; breast; squamous cell; melanoma;
 KW sarcoma; malignancy; VEGF-C; vascular endothelial growth factor C.
 XX
 OS Homo sapiens.
 XX
 XX
 XX Key
 FH Location/Qualifiers
 FH 1..31
 FT /label= Signal_peptide
 FT 32..103
 FT /label= N-terminal_peptide
 FT /note= "cleavage of this peptide from partially processed
 FT VEGF-C produces a fully processed mature form of VEGF-C
 FT of 21-23 KD which has high affinity to VEGFR-2"
 FT 104..227
 FT /label= Mature_VEGF-C
 FT 228..419
 FT /label= C-terminal_peptide
 FT /note= "Has a pattern of spaced cysteine residues
 FT reminiscent of a Balbiani ring 3 protein (BR3P) sequence;
 FT

[illegible]

XX	Human vascular endothelial growth factor (VEGF)-C protein.	
XX		
KW	Vascular endothelial growth factor-C; VEGF; human; re-endothelialisation;	
KM	vascular endothelial growth factor receptor; VEGFR; vascular trauma;	
KW	blood vessel; cardiovascular surgery; anti-restenosis agent; prevention;	
KM	restenosis; stenosis; percutaneous transluminal coronary angioplasty.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	1..31
FT		/label= "Signal peptide
FT		/note= "Cleavage results in partially-processed VEGF-C
FT		protein (29 kD)"
FT	Peptide	32..103
FT		/label= "Amino-terminal peptide
FT		/note= "Cleavage results in fully-processed mature
FT		VEGF-C protein (21-23 kD)"
FT	Protein	104..227
FT		/label= "Mature_human_VEGF_C
FT		/note= "Processed vascular epithelial growth factor-C"
FT	Binding-site	83
FT		/note= "Essential for VEGFR-2 and VEGFR-3 binding"
FT	Active-site	131..211
FT		/note= "Essential for biological activity of protein"
FT	Binding-site	137
FT		/note= "Essential for VEGFR-2 and VEGFR-3 binding"
FT	Binding-site	156
FT		/note= "Essential for VEGFR-2 binding"
FT	Binding-site	165
FT		/note= "Essential for VEGFR-2 and VEGFR-3 binding"
FT	Peptide	228..419
FT		/label= "Carboxy terminal peptide
FT		/note= "Cleavage results in partially-processed VEGF-C
FT		protein (29 kD)"
PN	WO200024412-A2.	
XX		
PD	04-MAY-2000.	
XX		
PF	26-OCT-1999; 99WO-US24054.	
XX		
PR	26-OCT-1998; 98US-0105587.	
XX		
XX	(LUDWIG INST CANCER RES.	
PA	(UYHE-) UNIV HELSINKI LICENSING LTD OY.	
PA	(YLAH/) YLA-HERTUULA S.	
XX		
PI	Yla-herthuula S, Alltalo K, Hiltunen MO, Jeltsch MM, Achen MG;	
XX		
DR	WPI; 2000-350584/30.	
XX		
DR	N-PSDB; AAD00339, AAD00353.	
XX		
PT	Preventing stenosis and restenosis in mammals using vascular	
PT	endothelial growth factor proteins or the nucleic acids encoding them -	
XX		
XX	Claim 5; Page 51-53; 61pp; English.	
XX		
PS		
CC	The present amino acid sequence is the complete human prepro-vascular	
CC	endothelial growth factor (VEGF)-C. VEGF-C has the ability to stimulate	
CC	re-endothelialisation of an injured blood vessel, without significant	
CC	stimulation of smooth muscle cell proliferation. It can bind to and	
CC	stimulate VEGFR-2 (vascular endothelial growth factor receptor) and/or	
CC	VEGF-3 phosphorylation in cells that express such receptors. An	
CC	anti-restenosis agent comprising either a VEGF-C gene or protein is	
CC	used in a method to reduce or prevent restenosis and stenosis of a blood	
CC	vessel following vascular trauma e.g., cardiovascular surgery and	
CC	percutaneous transluminal coronary angioplasty.	
XX		
XX		
Sequence	419 AA;	

Query Match 100.0%; Score 539; DB 21; Length 419;
 Best Local Similarity 100.0%; Pred. No. 1.1e-55;
 Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAFESGLDSDAEPDAGEATAYASKDLEQLRSVSSVDELMATVLYPEYWKMYKCOLRKG 60
 |||||||
 DB 30 aafesglidsdaepdagetaayaskdleeqllrsvssvdelmvtllypeywkmykcolrkxg 89

OY 61 WOHNRQANLNSRTEETIKFAAAHYNTETIKSIDNEMRKTC 102
 |||||||
 DB 90 wqhnrqanlnsrteetlkfaahyntetllksidnewrktgc 131

RESULT 15
 AAY97570
 ID AAY97570 standard; Protein; 419 AA.
 AC AAY97570;
 DT 05-APR-2001 (first entry)
 XX
 DE Human VEGF-B protein sequence.
 XX
 KW Human; angiogenic protein; wound healing; vascular tissue repair;
 KW peripheral arterial disease; critical limb ischaemia; coronary disease;
 KW angiogenesis; tumour; inflammation; diabetic retinopathy; psoriasis;
 KW rheumatoid arthritis; autoimmune disease; allergy; cancer; therapy;
 KW infectious disease; neurodegeneration;
 KW vascular endothelial growth factor-B; VEGF-B.
 KM
 XX
 OS Homo sapiens.
 OS
 XX
 PN WO200075163-A1.
 XX
 PD 14-DEC-2000.
 XX
 PF 01-JUN-2000; 2000WO-US14925.
 XX
 PR 03-JUN-1999; 99US-0137796.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Hu J, Cao L;
 PI
 DR WPI: 2001-071057/08.
 DR N-PSDB: AAA91004.
 XX
 PT New nucleic acid encoding angiogenic proteins, useful e.g. for
 PT promoting healing of wounds and treating peripheral arterial disease,
 PT critical limb ischaemia or coronary disease -
 XX
 PS Claim 11; Fig 1; 244pp; English.
 XX
 CC This sequence is vascular endothelial growth factor-B (VEGF-B),
 CC which is an angiogenic protein of the invention. The angiogenic proteins
 CC and the DNA sequences encoding them, are used to prevent, treat or
 CC ameliorate disease and to detect diseases, or susceptibility, by
 CC detecting mutations or the presence or amount of angiogenic protein
 CC expression. Particularly they are used to stimulate wound healing,
 CC growth of damaged bone and tissue, and for repair of vascular tissue,
 CC especially peripheral arterial disease, critical limb ischaemia or
 CC coronary disease. Antagonists of the sequences are used to inhibit
 CC angiogenesis in tumours and to treat inflammation (where associated with
 CC increased vascular permeability), diabetic retinopathy, Rheumatoid
 CC arthritis or psoriasis. Agonists are also useful for stimulating
 CC (lymph)angiogenesis. The proteins are also used to identify specific
 CC binding agents (potential therapeutic agents) and to raise antibodies.
 CC The antibodies are useful as therapeutic (antagonists; for detection,
 CC purification and targeting of proteins for in vivo or in vitro diagnosis
 CC (including imaging) or for therapy (including when linked to e.g. a label
 CC or cytotoxin); and for immunotyping of cells, e.g. for detecting minimal
 CC residual disease or haematopoietic progenitor/stem cells. It is also
 CC contemplated that the sequences might be useful for treating a very wide

CC range of other disorders, e.g. autoimmune diseases; allergy; cancer;
 CC infectious diseases (viral, bacterial, fungal or parasitic);
 CC neurodegeneration, also as chemotactic agents or for stimulating
 CC regeneration of the nervous system etc.
 CC
 XX

Sequence 419 AA;

Query Match 100.0%; Score 539; DB 22; Length 419;
 Best Local Similarity 100.0%; Pred. No. 1.1e-55;
 Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAFESGLDSDAEPDAGEATAYASKDLEQLRSVSSVDELMATVLYPEYWKMYKCOLRKG 60
 |||||||
 DB 30 aafesglidsdaepdagetaayaskdleeqllrsvssvdelmvtllypeywkmykcolrkxg 89

OY 61 WOHNRQANLNSRTEETIKFAAAHYNTETIKSIDNEMRKTC 102
 |||||||
 DB 90 wqhnrqanlnsrteetlkfaahyntetllksidnewrktgc 131

Search completed: October 17, 2001, 14:45:59
 Job time: 125 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:47:08 ; Search time 62.93 Seconds
(without alignments)
33.374 Million cell updates/sec

Title: US-09-427-657-2_COPY_30_131
Perfect score: 539
Sequence: 1 AAFESGIDLSDAEPDAGEAT.....AHYNTFLIKSIDNEWKRTQC 102

Scoring table: BIOSUM62
Gapop 10.0 ; Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued Patents-AA:
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCUTS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	539	100.0	419	2	US-08-999-811-2
2	539	100.0	419	3	US-09-042-105-2
3	539	100.0	419	3	US-09-042-105-18
4	539	100.0	419	4	US-08-795-430-8
5	539	100.0	419	4	US-08-510-133A-35
6	533	98.9	419	5	PCT-US96-09001-2
7	425	78.8	415	4	US-08-795-430-11
8	412	76.4	418	4	US-08-795-430-13
9	347	64.4	350	2	US-08-999-811-4
10	347	64.4	350	2	US-08-824-996-2
11	347	64.4	350	3	US-09-042-105-4
12	347	64.4	350	4	US-08-510-133A-33
13	347	64.4	350	4	US-08-585-895-33
14	163	30.2	321	4	US-08-915-795-9
15	163	30.2	358	4	US-08-915-795-8
16	155	28.8	325	4	US-08-915-795-3
17	155	28.8	354	4	US-08-915-795-5
18	86	16.0	17	4	US-08-795-430-9
19	86	16.0	18	4	US-08-795-430-5
20	86	16.0	18	4	US-08-510-133A-13
21	69.5	12.9	109	2	US-08-585-895-13
22	69.5	12.9	110	3	US-08-070-116A-4
23	69.5	12.9	326	3	US-08-444-644-44
24	69.5	12.9	327	2	US-08-808-720-3
25	69.5	12.9	328	3	US-08-761-277A-47
26	69.5	12.9	331	3	US-08-808-720-1
27	69.5	12.9	331	3	US-08-808-720-5

28	69.5	12.9	331	3	US-08-808-720-7	Sequence 7, Appl
29	69.5	12.9	382	1	US-08-470-299-7	Sequence 7, Appl
30	69.5	12.9	382	1	US-08-470-299-10	Sequence 10, Appl
31	69.5	12.9	443	5	PCT-US96-13152-4	Sequence 4, Appl
32	69.5	12.9	467	1	US-07-704-744-81	Sequence 8, Appl
33	69.5	12.9	467	2	US-07-916-098A-45	Sequence 45, Appl
34	69.5	12.9	467	4	US-08-523-894-8	Sequence 8, Appl
35	69.5	12.9	467	4	US-08-523-894-10	Sequence 10, Appl
36	69.5	12.9	467	4	US-08-523-894-12	Sequence 12, Appl
37	65	12.1	117	1	US-07-614-443A-2	Sequence 2, Appl
38	65	12.1	117	1	US-08-294-859-2	Sequence 2, Appl
39	65	12.1	117	1	US-08-481-676-2	Sequence 2, Appl
40	64.5	12.0	148	1	US-08-686-178-2	Sequence 2, Appl
41	64.5	12.0	3324	2	US-08-705-660-34	Sequence 34, Appl
42	64.5	12.0	3324	2	US-08-989-045-34	Sequence 34, Appl
43	64	11.9	238	5	PCT-US94-01149-24	Sequence 24, Appl
44	63.5	11.8	657	3	US-08-893-852A-3	Sequence 3, Appl
45	63.5	11.8	657	4	US-08-821-818-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-999-811-2
; Sequence 2, Application US/08999811
; Patent No. 5932540
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: CAO, LIANG
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/999, 811
; FILING DATE: HEREWITH
; CLASSIFICATION:
; APPLICATION NUMBER: US 08/207, 550
; FILING DATE: 8-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465, 968
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: MARKOWITZ, KAREN R.
; REGISTRATION NUMBER: 36,351
; REFERENCE/DOCKET NUMBER: 1488.1000004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO. 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-999-811-2
Query Match 100.0%; Score 539; DB 2; Length 419;
Best Local Similarity 100.0%; Pred. No. 1e-60;

Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAFESGLDSDAPDAGEATAYASKDLEQLRSVSSVDLMTLVLPYWKMYKCOLRKG 60
|||||
DB 30 AAFESGLDSDAPDAGEATAYASKDLEQLRSVSSVDLMTLVLPYWKMYKCOLRKG 89

QY 61 WOHNRQANLNSRTEETIKFAAHYNTLKSIDNEMRKTQC 102
|||||
DB 90 WOHNRQANLNSRTEETIKFAAHYNTLKSIDNEMRKTQC 131

RESULT 2
US-09-042-105-2
; Sequence 2, Application US/09042105
; Patent No. 6040157
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: CAO, LIANG
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,105
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/207,550
; FILING DATE: 8-MAR-1994
; CLASSIFICATION:
; APPLICATION NUMBER: US 08/465,968
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TO BE ASSIGNED
; FILING DATE: 24-DEC-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: ERIC K. STEFFE
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-042-105-2

Query Match 100.0%; Score 539; DB 3; Length 419;
Best Local Similarity 100.0%; Pred. No. 1e-60;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAFESGLDSDAPDAGEATAYASKDLEQLRSVSSVDLMTLVLPYWKMYKCOLRKG 60
|||||
DB 30 AAFESGLDSDAPDAGEATAYASKDLEQLRSVSSVDLMTLVLPYWKMYKCOLRKG 89

QY 61 WOHNRQANLNSRTEETIKFAAHYNTLKSIDNEMRKTQC 102
|||||
DB 90 WOHNRQANLNSRTEETIKFAAHYNTLKSIDNEMRKTQC 131

RESULT 3
US-09-042-105-18
; Sequence 18, Application US/09042105
; Patent No. 6040157
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: CAO, LIANG
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,105
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,968
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; APPLICATION NUMBER: US 08/207,550
; FILING DATE: 8-MAR-1994
; CLASSIFICATION:
; APPLICATION NUMBER: TO BE ASSIGNED
; FILING DATE: 24-DEC-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: ERIC K. STEFFE
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-042-105-18

Query Match 100.0%; Score 539; DB 3; Length 419;
Best Local Similarity 100.0%; Pred. No. 1e-60;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAFESGLDSDAPDAGEATAYASKDLEQLRSVSSVDLMTLVLPYWKMYKCOLRKG 60
|||||
DB 30 AAFESGLDSDAPDAGEATAYASKDLEQLRSVSSVDLMTLVLPYWKMYKCOLRKG 89

QY 61 WOHNRQANLNSRTEETIKFAAHYNTLKSIDNEMRKTQC 102
|||||
DB 90 WOHNRQANLNSRTEETIKFAAHYNTLKSIDNEMRKTQC 131

```
RESULT 4
US-08-795-430-8
; Sequence 8, Application US/08795430
; Patent No. 6130071
; GENERAL INFORMATION:
; APPLICANT: Aitalo, Kari
; APPLICANT: Joukov, Vladimir
; TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
; TITLE OF INVENTION: Protein and Gene, Mutants Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,430
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/E196/00427
; FILING DATE: 01-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/671,573
; FILING DATE: 28-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/601,132
; FILING DATE: 14-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/585,895
; FILING DATE: 12-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/510,133
; FILING DATE: 01-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,011
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 28967/33691
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-795-430-8

Query Match          100.0%; Score 539; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 1e-60;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
RESULT 5
US-08-510-133A-35
; Sequence 35, Application US/08510133A
; Patent No. 6221839
; GENERAL INFORMATION:
; APPLICANT: Aitalo, Kari
; APPLICANT: Joukov, Vladimir
; TITLE OF INVENTION: Receptor Ligand
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/510,133A
; FILING DATE: 01-AUG-1995
; CLASSIFICATION: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 28113/32863
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-08-510-133A-35

Query Match          100.0%; Score 539; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 1e-60;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
RESULT 6
PCT-US96-09001-2
; Sequence 2, Application PC/TUS9609001
; GENERAL INFORMATION:
; APPLICANT: HU, ET AL.
; TITLE OF INVENTION: Human Vascular Endothelial Growth Factor 2
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILLILLAN,
; ADDRESS: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
```

```

MEDUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/09001
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/465,968
FILING DATE: 6 JUN 95
APPLICATION NUMBER: 08/207,550
FILING DATE: 8 MAR 1994
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-288
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
PCT-US96-09001-2

```

Query Match	98.9%;	Score 533;	DB 5;	Length 419;
Best Local Similarity	99.0%;	Pred. No. 5.9e-60;		
Matches 101; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

Qy	1	AAESGSDLSDAEPDGGETAATAAKDLSEORSSVDELMTVLYPEYMKYCOLRKG	60
Db	30	AAESGSDLSDAEPDGGETAATAAKDLSEORSSVDELMTVLYPEYMKYCOLRKG	89
Qy	61	WOHNRQANLNSTRETTKFAAAHYNTIELKSIDNEMKRTQC	102
Db	90	WOHNRQANLNSTRETTKFAAAHYNTIELKSIDNEMKRTQC	131

RESULT 7
US-08-795-430-11
; Sequence 11, Application US/08795430

APPLICANT: Altalo, Kari
 APPLICANT: Joukov, Vladimir
 TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
 TITLE OF INVENTION: Protein and Gene, Mutants Thereof, and Uses Thereof
 NUMBER OF SEQUENCES: 57
 CORRESPONDENCE ADDRESS:

ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive

STATE: Illinois
COUNTRY: United States of America

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.30

APPLICATION NUMBER: US/08/795,430

CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCI/FI98/0042/
FILING DATE: 01-AUG-1996

PRIOR APPLICATION DATA:

```

1      APPLICATION NUMBER: 08/671,573
2      FILING DATE: 28-JUN-1996
3      PRIOR APPLICATION DATA:
4      APPLICATION NUMBER: 08/601,132
5      FILING DATE: 14-FEB-1996
6      PRIOR APPLICATION DATA:
7      APPLICATION NUMBER: 08/585,895
8      FILING DATE: 12-JUN-1996
9      PRIOR APPLICATION DATA:
10     APPLICATION NUMBER: 08/510,133
11     FILING DATE: 01-AUG-1995
12     PRIOR APPLICATION DATA:
13     APPLICATION NUMBER: 08/340,011
14     FILING DATE: 14-NOV-1994
15     ATTORNEY/AGENT INFORMATION:
16     NAME: GASS, David A.
17     REGISTRATION NUMBER: 38,153
18     REFERENCE/DOCKET NUMBER: 28967/336991
19     TELECOMMUNICATION INFORMATION:
20     TELEPHONE: 312/474-6300
21     TELEFAX: 312/474-0448
22     TELEX: 25-3856
23     INFORMATION FOR SEQ ID NO: 11:
24     SEQUENCE CHARACTERISTICS:
25     LENGTH: 415 amino acids
26     TYPE: amino acid
27     TOPOLOGY: linear
28     MOLECULE TYPE: protein
29     OS-08-795-430-11

```

Query Match	78.8%;	Score 425;	DB 4;	Length 415;
Best Local Similarity	78.4%;	Pred. No. 3.8e-46;		
Matches 80;	Conservative 8;	Mismatches 10;	Indels 4;	Gaps 1

QY 1 AAFESGLDSDAEPDAGEATATAYASKDLEEDLRSSVSDLMVLYPEYKMYKQCLRKG 600

Db 30 AAFESGLGFEAEPPDGGEVKAFEGKDLIEDLRSSVSDLMVLYPYDKMYKQCLRKG 899

```
QY      61 WQHNRQANLSRTEETIKFAAAHYNTEILKSIDNEWKRQTQC 1022
        || - |::|| ::||| ||||| ||||| |||||
Db      90 WQ-----QPTLNTRTGDGVKFAAAHYNTEILKSIDNEWKRQTQC 1277
```

RESULT 8
US-08-795-430-13

; Sequence 13, Application US/087954330
; Patent No. 6130071

APPLICANT: Altalo, Kari
 APPLICANT: Joukov, Vladimir
 TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
 TITLE OF INVENTION: Protein and Gene, Mutants Thereof, and Uses Thereof
 NUMBER OF SEQUENCES: 57
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago

COUNTRY: United States of America

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

APPLICATION NUMBER: US/08/795,430
 FILING DATE: .

CLASSIFICATION: 433
PRIOR APPLICATION DATA:

AFFIDAVIT NUMBER: FC1/F130/0042
 FILING DATE: 01-AUG-1996

MOLECULE TYPE: protein
US-08-795-430-11

Patent No. 6040157
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,105
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550
FILING DATE: 8-MAR-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TO BE ASSIGNED
FILING DATE: 24-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: ERIC K. STEFFE
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-042-105-4

Query Match 64.4%; Score 347; DB 3; Length 350;
Best Local Similarity 100.0%; Pred. No. 2.8e-36;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 41 MVLVPEYWKMKCOLRKGWQHNRQANLSRTEETIKFAAAHYNTLKSINDENWRT 100
DB 1 MVLVPEYWKMKCOLRKGWQHNRQANLSRTEETIKFAAAHYNTLKSINDENWRT 60
QY 101 QC 102
DB 61 QC 62
RESULT 12
US-08-510-133A-33
Sequence 33, Application US/08510133A
Patent No. 6221839
GENERAL INFORMATION:
APPLICANT: Allitalo, Kari
Joukov, Vladimir
TITLE OF INVENTION: Receptor Ligand
NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/510,133A
FILING DATE: 01-Aug-1995
CLASSIFICATION: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28113/32863
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-08-510-133A-33

Query Match 64.4%; Score 347; DB 4; Length 350;
Best Local Similarity 100.0%; Pred. No. 2.8e-36;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 41 MVLVPEYWKMKCOLRKGWQHNRQANLSRTEETIKFAAAHYNTLKSINDENWRT 100
DB 1 MVLVPEYWKMKCOLRKGWQHNRQANLSRTEETIKFAAAHYNTLKSINDENWRT 60
QY 101 QC 102
DB 61 QC 62
RESULT 13
US-08-585-895-33
Sequence 33, Application US/08585895
Patent No. 6245530
GENERAL INFORMATION:
APPLICANT: Joukov, Vladimir
TITLE OF INVENTION: Receptor Ligand
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/585,895
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:

NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28113/33072
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-585-895-33

Query Match 64.4%; Score 347; DB 4; Length 350;
Best Local Similarity 100.0%; Pred. No. 2.8e-36;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 41 MTVLPEYKMKYKCOLRKGMQHNREOANLNSRTETIKFAAHYNTILKSIDNEMKRT 100
DB 1 MTVLPEYKMKYKCOLRKGMQHNREOANLNSRTETIKFAAHYNTILKSIDNEMKRT 60

OY 101 QC 102
DB 61 QC 62

RESULT 14
US-08-915-795-9
Sequence 9, Application US/08915795
Patent No. 6235713

GENERAL INFORMATION:

APPLICANT: Marc G. ACHEN
APPLICANT: Andrew F. WILKS
APPLICANT: Steven A. STACKER
APPLICANT: Kari ALITALO
TITLE OF INVENTION: GROWTH FACTOR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKee, Edwards & Lenahan P.L.L.C.
STREET: 1200 G Street, NW, Suite 700
CITY: Washington
STATE: DC
COUNTRY: United States of America
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,795
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/42983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
TELEX: N/A

INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE: Mouse Lung
TISSUE TYPE: Mouse Lung

US-08-915-795-9

Query Match 30.2%; Score 163; DB 4; Length 321;
Best Local Similarity 39.0%; Pred. No. 8.6e-13;
Matches 30; Conservative 23; Mismatches 16; Indels 8; Gaps 2;

OY 28 EEOILSVSSVDELMVLPYKMKYKCOLRKGMQHNREOANLNSR--TEETIKFAAHY 85
DB 41 EEOILSVSSVDELMVLPYKMKYKCOLRKGMQHNREOANLNSR--TEETIKFAAHY 94

OY 86 NTEILKSIDNEMKRTQC 102
DB 95 DTEILKVIDEEMWORTQC 111

RESULT 15
US-08-915-795-8
Sequence 8, Application US/08915795
Patent No. 6235713

GENERAL INFORMATION:
APPLICANT: Marc G. ACHEN
APPLICANT: Andrew F. WILKS
APPLICANT: Steven A. STACKER
APPLICANT: Kari ALITALO
TITLE OF INVENTION: GROWTH FACTOR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKee, Edwards & Lenahan P.L.L.C.
STREET: 1200 G Street, NW, Suite 700
CITY: Washington
STATE: DC
COUNTRY: United States of America
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,795
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/42983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
TELEX: N/A

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 358 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE: Mouse Lung
TISSUE TYPE: Mouse Lung
US-08-915-795-8

Query Match 30.2%; Score 163; DB 4; Length 358;
Best Local Similarity 39.0%; Pred. No. 1e-12;
Matches 30; Conservative 23; Mismatches 16; Indels 8; Gaps 2;

OY 28 EEOILSVSSVDELMVLPYKMKYKCOLRKGMQHNREOANLNSR--TEETIKFAAHY 85
DB 46 EEOILSVSSVDELMVLPYKMKYKCOLRKGMQHNREOANLNSR--TEETIKFAAHY 99
OY 86 NTEILKSIDNEMKRTQC 102

Wed Oct 17 15:32:33 2001

us-09-427-657-2_copy_30_131.ra1

Page 8

Db 100 DFTLKVDEWQRTQC 116

Search completed: October 17, 2001, 14:47:09
Job time: 195 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:48:34 ; Search time 78.16 Seconds
(without alignments)
99.409 Million cell updates/sec

Title: US-09-427-657-2_COPY_30_131

Perfect score: 539
Sequence: 1 AAFESGLDLSDAEPDAGEAT.....AHYNTETLKSIDNWRKTC 102

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	539	100.0	419	2	S69207
2	77	14.3	1494	2	T14355
3	70.5	13.1	1729	2	T43403
4	70	13.0	234	2	I64046
5	70	13.0	371	2	A44122
6	69.5	12.9	327	1	GAHU
7	69.5	12.9	448	2	A69959
8	69	12.8	786	2	T16509
9	69	12.8	1403	2	T49093
10	69	12.8	4644	1	A38905
11	68.5	12.7	512	2	S70644
12	68.5	12.7	984	2	F81299
13	68.5	12.7	1265	2	S57968
14	68	12.6	489	2	T27468
15	68	12.6	611	2	S52229
16	67	12.4	329	2	T03059
17	67	12.4	329	2	T03688
18	67	12.4	375	2	C71441
19	67	12.4	453	2	H81151
20	67	12.4	453	2	D81870
21	67	12.4	493	2	T29030
22	67	12.4	532	2	E75617
23	66	12.4	703	2	B82148
24	66	12.2	251	2	F72114
25	66	12.2	251	2	A86508
26	66	12.2	1500	1	J01348
27	65.5	12.2	532	2	T02539
28	65	12.1	239	2	E75445
29	65	12.1	246	2	T14772

30	65	12.1	503	1	YF8YAC	phenylalanine--trn
31	65	12.1	2408	2	T24483	hypothetical prote
32	65	12.1	2413	2	S34670	hypothetical prote
33	64.5	12.0	363	2	F65070	hypothetical prote
34	64.5	12.0	396	2	C85942	probable carbamoyl
35	64.5	12.0	437	2	I51238	translational elonga
36	64.5	12.0	442	2	G84465	probable cytochrom
37	64.5	12.0	479	2	S29301	alliin lyase (EC 4
38	64.5	12.0	485	2	T25199	hypothetical prote
39	64.5	12.0	712	2	E81196	transferrin-bindin
40	64.5	12.0	843	2	D96495	unknown protein (l
41	64.5	12.0	3224	1	S58884	Ran-binding protei
42	64	11.9	284	2	A75152	cysteine synthase
43	64	11.9	1325	1	S73723	probable lipoprote
44	64	11.9	1449	2	B84426	hypothetical prote
45	63.5	11.8	120	2	E71980	hypothetical prote

ALIGNMENTS

RESULT 1
S69207
vascular endothelial growth factor C precursor - human
N:Alternate names: FLT4 ligand DHM
C:Species: Homo sapiens (man)
C>Date: 27-Apr-1996 #sequence, revision 01-Nov-1996 #text_change 08-Oct-1999
C:Accession: S69207; S61795; S71443; S69208; G02659
R:Joukov, V.; Pajusola, K.; Kaipainen, A.; Chillov, D.; Lahtinen, I.; Kukk, E.; Saksel
EMBO J. 15, 1751, 1996
A>Title: Corrigendum: A novel vascular endothelial growth factor, VEGF-C, is a ligand
A:Reference number: S69207; MUID:96203094
A:Accession: S69207
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-419 <JOU>
A:Cross-references: EMBL:X94216; NID:g1177488; PIDN:CAA63907.1; PID:e221096; PID:g118
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1995
A>Note: only a part of the translation is shown
A>Note: this is a revision to the sequence from reference S61795
R:Joukov, V.; Pajusola, K.; Kaipainen, A.; Chillov, D.; Lahtinen, I.; Kukk, E.; Saksel
EMBO J. 15, 290-298, 1996
A>Title: A novel vascular endothelial growth factor, VEGF-C, is a ligand for the Flt4
A:Reference number: S61795; MUID:96178224
A:Accession: S61795
A>Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 70-419 <JOU1>
A>Note: this sequence has been revised in reference S69207
A:Accession: S71443
A:Molecule type: protein
A:Residues: X, 104-120 <JOU2>
A:Status: X, 104-120 <JOU2>
R:Lee, J.; Gray, A.; Yuan, J.; Luoh, S.M.; Avraham, H.; Wood, W.I.
submitted to the EMBL Data Library, December 1995
A:Description: Vascular endothelial growth factor related protein (VRP): A ligand and
A:Reference number: S69208
A:Accession: S69208
A:Molecule type: mRNA
A:Residues: 1-419 <LEB>
A:Cross-references: EMBL:U43142; NID:g1150988; PIDN:AAA85214.1; PID:g1150989
R:Morris, J.C.
submitted to the EMBL Data Library, May 1996
A:Reference number: H01557
A:Accession: G02655
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-419 <MOR>
A:Cross-references: EMBL:U58111; NID:g1373426; PIDN:AAB02909.1; PID:g1373427
C:Genetics:
A:Gene: GDB:VEGFC; VRP
A:Cross-references: GDB:3890883; OMIM:601528
F:1-12/Domain: signal sequence #status predicted <SIG>
F:13-102/Domain: propeptide #status predicted <PRO>

A>Note: sequence extracted from NCBI backbone (NCBIN:106718, NCBIPI:106719)

Query Match 13.0%; Score 70; DB 2; Length 371;

Best Local Similarity 31.2%; Pred. No. 11;

Matches 24; Conservative 12; Mismatches 27; Indels 14; Gaps 3;

QY 16 AGATATYASKDLEEQLRSSVDELMTVLYPEYKWK-----YKCOLKRGQWQHNRQANL 70

DB 109 AGEQTQARRELADAAQVDDLEFKDELQDRYDLSLTALLESERSLQ--QHDVEMAO 165

QY 71 NSRTEETIKFAAAHYNT 87

DB 166 KER-----LAAAEENT 176

RESULT 6
G4HU
Ig gamma-4 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 16-Jul-1999

C:Accession: A90933; A90249; A02150

R:Ellison, J.; Buxbaum, J.; Hood, L.

DNA 1, 11-18, 1981

A:Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.

A:Reference number: A90933; MUID:83157104

A:Accession: A90933

A:Molecule type: DNA

A:Residues: 1-327 <ELL>

A>Note: the sequence was determined from the germline gene

R:Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.

Biochem. J. 117, 33-47, 1970

A:Title: Human immunoglobulin subclasses. Partial amino acid sequence of the constant

A:Accession: A90249

A:Molecule type: Protein

A:Residues: 1-30;81-326 <PIN>

C:Genetics:

A:Gene: GDB:IGHG4

A:Cross-references: GDB:119340; OMIM:147130

A:Map position: 14q32.33-14q32.33

A:Introns: 99/1; 111/1; 221/1

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap

hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1a

C:Superfamily: Immunoglobulin C region; immunoglobulin homology

C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin

F:20-85/Domain: immunoglobulin homology <IM1>

F:134-203/Domain: immunoglobulin homology <IM2>

F:240-307/Domain: immunoglobulin homology <IM3>

F:14/Disulfide bonds: interchain (to light chain) #status experimental

F:27-83,141-201,247-305/Disulfide bonds: #status predicted

F:106-109/Disulfide bonds: interchain (to heavy chain) #status experimental

F:177/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 12.9%; Score 69.5; DB 1; Length 327;

Best Local Similarity 29.4%; Pred. No. 11;

Matches 25; Conservative 12; Mismatches 29; Indels 19; Gaps 3;

QY 7 IDLSDAEP-----DAGEATAYASKDLEEQLRSSVDELMTVLYPEY--WKMYKCOL 56

DB 144 VVDSQDEDFEVNVDGQEVNHNKTKPREDEGFNTYRVSVLTVLHDDMLNGKRYKKCV 203

QY 57 RKGGMQHNREQANLNSRTEETIKFA 81

DB 204 SNKG-----LPSSEIKETISK 219

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C:Accession: A69959

R:Kunst, F.; Ogasawara, N.; Mosser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber

C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.;

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,

Nature 390, 249-256, 1997

A:Authors: Foulger, C.R.; Frits, C.; Fujita, M.; Fujita, Y.; Funai, S.; Galizzi, A.; Gal

isch, J.; Hardwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M

koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete

Rieger, M.; Rivaletta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanl

A:Authors: Schleich, S.; Schoeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Togonni, A.; Tosato, V.; Uchiya

T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*

A:Reference number: A69580; MUID:98044033

A:Accession: A69959

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-448 <KUN>

A:Cross-references: GB:299116; GB:AL009126; NID:92634723; PIDN:CAB14387.1; PID:926348

A:Experimental source: strain 168

C:Genetics:

A:Gene: yqhJ

C:Superfamily: Aquifex aeolicus glycine dehydrogenase (decarboxylating)

Query Match 12.9%; Score 69.5; DB 2; Length 448;

Best Local Similarity 28.4%; Pred. No. 16;

Matches 23; Conservative 17; Mismatches 28; Indels 13; Gaps 4;

QY 25 KLEEQLRNS--VSSVDELMTVLYPE---YKMYKCOLKRGQWQHNRQANLNSRTEETIK 79

DB 11 KDKQEMLATIGVSSIDDLFADI-PENWYKXKREHOKAKSETELRELTKLASKNRDVO 69

QY 80 FAA-----AHYNTETIKSI 93

DB 70 YASFLGAGYDHYKQYVDHV 90

RESULT 8

T16509

hypothetical protein F59A6.3 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T16509

R:Nhan, M.

submitted to the EMBL Data Library, December 1995

A:Description: The sequence of C. elegans cosmid F59A6.

A:Reference number: Z18526

A:Accession: T16509

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-786 <NHA>

A:Cross-references: EMBL:U41994; NID:g1123047; PID:g1123052; PIDN:AA83456.1; CESP:F5

A:Introns: 106/1; 620/1; 634/3; 675/3; 775/1

Query Match 12.8%; Score 69; DB 2; Length 786;

Best Local Similarity 33.3%; Pred. No. 35;

Matches 18; Conservative 9; Mismatches 17; Indels 10; Gaps 2;

QY 22 YASKDLEQLRSVSSVDELMTVLYPEY-----RMKCOLKRGQWQHNR 66

DB 719 YTSPTSSQIKSSYSGELITQVCPQSYVFETALQPKAIYKC-LKTGMWAGSPE 771

RESULT 9
A69959
glycine dehydrogenase homolog yqhJ - *Bacillus subtilis*

C:Species: *Bacillus subtilis*

Db 133 DIKSDDEKGAIIKAIKONEKLANVQSVGE---YLKEYFOKFK-----ENSKER 181
QY 68 ANNSRTEETIKFAAAHYNTIELKSIDNEWRK 99
Db 182 TNVNRKSEYERCIQSFLEKELIKFKORE 213

RESULT 13

S57968
Ran-binding protein 2 - mouse (fragment)
N:Alternate names: RanBP2 protein
C:Species: Mus musculus (house mouse)
C>Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 05-Nov-1999
C:Accession: S57968
R:Wilken, N.; Senecal, J.L.; Scheer, U.; Dabaivalle, M.C.
submitted to the EMBL Data Library, May 1995
A:Description: Localization of the Ran-GTP binding protein RanBP2 at the cytoplasmic side
A:Reference number: S57968
A:Accession: S57968
A:Molecule type: mRNA
A:Residues: 1-1265 <MIL>
A:Cross-references: EMBL:X87337; NID:g8899335; PIDN:CAA60778.1; PID:g1065884
C:Genetics:
A:Gene: RanBP2
C:Superfamily: nucleoporin Nup358; cyclophilin homology; tetratricopeptide repeat homolo

Query Match 12.7%; Score 68.5; DB 2; Length 1265;
Best Local Similarity 19.6%; Pred. No. 68;
Matches 27; Conservative 21; Mismatches 53; Indels 37; Gaps 3;

QY 2 AFESGLDSDAEPDAGEATAYASKDLEE-----OLRSVSYDELM 41
Db 1123 SFSALSPSPKSPAKLNQSGASVCTDESDVTOEERDQGFEPVLPDLVEVSSGEENE 1182
QY 42 TVLYPEYKMYKCOLRKGQ-----HNREQANLSRTEETIKFAAAHYNT-- 87
Db 1183 QVFNHRAKLYRDKDVGQWKEGIGDIKILQVNDKQVRIVMRQVILKICANHRITPD 1242
QY 88 ---ELKSIDNEWRKTQC 102
Db 1243 MTLQTMKGTERTVWWTAC 1260

RESULT 14

T27468
hypothetical protein Y87G2A.P - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C:Accession: T27468
R:White, S.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z20371
A:Accession: T27468
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-489 <MIL>
A:Cross-references: EMBL:AL110500; PIDN:CAB54490.1; CESP:Y87G2A.P
A:Experimental source: clone Y87G2A
C:Genetics:
A:Gene: CESP:Y87G2A.P
A:introns: 198/1; 265/3; 407/3

Query Match 12.6%; Score 68; DB 2; Length 489;
Best Local Similarity 29.8%; Pred. No. 25;
Matches 31; Conservative 13; Mismatches 36; Indels 24; Gaps 5;

QY 4 ESLGLDSDAEPDAGEATAYASKDLEQLRSVSYV---DELMTVLYPEYKMYK--COL-- 56
Db 61 ESDSDISDE---EATPPQMKELDKLVALLKEVFAAKVTETDPVKKRYNNLCQLST 116
QY 57 -----RKGMQHNRQANLSRTEETIKFAAAHYNTIELKSI 93

Db 117 TLKTISSVKGSGWK-----ILHKEHTEKELALDRANFDILKTV 154

RESULT 15

S52229
hypothetical protein 611 - Coxiella burnetii plasmids
C:Species: Coxiella burnetii
C>Date: 05-May-1995 #sequence_revision 01-Sep-1995 #text_change 08-Oct-1999
C:Accession: S52229; S38224
R:Williams, H.; Thiele, D.; Valkova, D.
submitted to the EMBL Data Library, March 1995
A:Reference number: S52227
A:Accession: S52229
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-611 <MIL>
A:Cross-references: EMBL:X85964; NID:g757760; PIDN:CAA59942.1; PID:g757763
A:Experimental source: strain Q1182, plasmid QPDV
R:Thiele, D.; Williams, H.; Haas, M.; Krauss, H.
submitted to the EMBL Data Library, October 1993
A:Reference number: S38215
A:Accession: S38224
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-140,'Q',142-155,'R',157-159,'T',161-170,'S',172-493,'S',495-530,'G',532
A:Cross-references: EMBL:X75356; NID:g407370; PIDN:CAA53112.1; PID:g407380
A:Experimental source: Isolate Nine Mile phase I, plasmid QPH1

Query Match 12.6%; Score 68; DB 2; Length 611;
Best Local Similarity 21.6%; Pred. No. 33;
Matches 24; Conservative 22; Mismatches 47; Indels 18; Gaps 3;

QY 7 LDISDAEP-----DAGEATAYASKDLEQLRSVSYVDELMTVLYPEYKMYKCOLR 58
Db 283 LSVEDKPAEIAIYLSREDAIATMTSEVLROKDPKLTITQELMSTPESVDRALTOLOK 342
QY 59 GGMQHNRQANLSRTEETIKFAAAHY-----NTEILSIDNEWRKT 100
Db 343 NGWVIESSVDRKGPISDSVQIAEHYCALSLERTNT--VIAAATNENROT 392

Search completed: October 17, 2001, 14:48:36
Job time: 282 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:51:38 ; Search time 42.92 Seconds

(without alignments)
81.409 Million cell updates/sec

Title: US-09-427-657-2_COPY_30_131

Perfect score: 539
Sequence: 1 AAFESGLDLDADPDAGEAT.....AHYNTLLKSIDNEMKRTQC 102

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : SWISSProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	539	100.0	419	1	VEGC_HUMAN
2	425	78.8	415	2	VEGC_MOUSE
3	70	13.0	198	1	GRPE_HAEN
4	69.5	12.9	327	1	GC4_HUMAN
5	69.5	12.9	448	1	GC51_BACSU
6	69	12.8	4644	1	DYHC_MOUSE
7	69	12.8	4644	1	DYHC_MOUSE
8	68.5	12.7	512	1	ANX7_XENLA
9	67	12.4	329	1	RIR2_TOBAC
10	66	12.2	251	1	Y137_CHLUP
11	66	12.2	1500	1	CPSM_HUMAN
12	65	12.1	502	1	SIFA_YEAST
13	65	12.1	2413	1	PR08_YEAST
14	64.5	12.0	347	1	FOS_CYPCA
15	64.5	12.0	363	1	YGEH_ECOCI
16	64.5	12.0	437	1	EPH1_XENLA
17	64.5	12.0	479	1	ALIN_ALICE
18	64.5	12.0	712	1	TPPB_NEIMB
19	64.5	12.0	3224	1	RBP2_HUMAN
20	64	11.9	1325	1	Y309_MYCPN
21	63.5	11.8	295	1	EAEH_ECOCI
22	63.5	11.8	441	1	PHPA_PLACH
23	63.5	11.8	657	1	MY16_MOUSE
24	63.5	11.8	664	1	STGB_AQUAE
25	63.5	11.8	1085	1	RBP2_BOVIN
26	63	11.7	619	1	OM70_NEUCR
27	63	11.7	997	1	YPX2_CAEEL
28	63	11.7	1272	1	Y228_METJA
29	62.5	11.6	436	1	EP1G_XENLA
30	62.5	11.6	447	1	ALIN_ALICE
31	62.5	11.6	447	1	ALIN_ALICE
32	62.5	11.6	640	1	Y6J9_YEAST
33	62.5	11.6	761	1	VPA_BPP2

34	62.5	11.6	767	1	RIR1_SYNY3	P74240 synchocyst
35	62.5	11.6	775	1	VPA_ROTAM	P11197 human rotav
36	62.5	11.6	1040	1	BO12_YEAST	P33969 saccharomyc
37	62	11.5	440	1	NAM1_YEAST	P10849 saccharomyc
38	61.5	11.4	123	1	ECHEB_ECHCA	P81996 echis carin
39	61.5	11.4	281	1	ATPG_PROMO	P29710 propionigen
40	61.5	11.4	394	1	MEMB_METTR	P27354 methylosinu
41	61.5	11.4	465	1	CORO_AICCA	O43989 actinobact
42	61.5	11.4	485	1	HDB_CLOAL	O96785 clogmia alb
43	61.5	11.4	740	1	GAG_SMRVH	P21411 squitrel mo
44	61.5	11.4	955	1	VP2_BPV17	P05309 bluetongue
45	61	11.3	76	1	CHAB_ECOCI	P39162 escherichia

ALIGNMENTS

RESULT	ID	VEGC_HUMAN	STANDARD	PRT	419 AA.
AC	P49767	VEGC_HUMAN	STANDARD	PRT	419 AA.
DT	01-OCT-1996	(Rel. 34, Created)			
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	01-OCT-2000	(Rel. 40, Last annotation, update)			
DE	VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR (VEGF-C) (VASCULAR				
DE	ENDOTHELIAL GROWTH FACTOR RELATED PROTEIN) (VRF) (FLT4 LIGAND) (FLT4-				
DE	L).				
GN	VEGFC.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
OX	NCBI_Taxid=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 103-120.				
RX	MEDLINE=96178224; PubMed=8617204;				
RA	Joukov V., Pejussola K., Kaipainen A., Chilov D., Lahtinen I., Kuk E.,				
RA	Saksela O., Kalkkinen N., Alitalo K.;				
RL	EMBO J. 15:1751-1751(1996).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=96312526; PubMed=8700872;				
RA	Lee J., Gray A., Yuan J., Luoh S.-M., Avraham H., Wood W.I.;				
RT	"Vascular endothelial growth factor-related protein: a ligand and				
RT	specific activator of the tyrosine kinase receptor Flt4.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 93:1988-1992(1996).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	Fitz L., Morris J.C., Towler P.S., Long A.J., Greco R.,				
RA	Burgess P., Giannotti J., Charletta A., Hennessey D., Kovacic S.,				
RA	Zillgerald M., Scaltrito H., Welch N., Neben S., Finerly H.,				
RA	Foltergerald M., Wang J., Nickbarg E., Gassaway R., Turner K.,				
RA	Wood C.R.;				
RL	Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.				
CC	-!- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL				
CC	CELL GROWTH.				
CC	-!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.				
CC	-!- PTM: PROBABLY PROTEOLITICALLY PROCESSED IN THE C-TERMINUS.				
CC	-!- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				

```

CC or send an email to license@lsb-sib.ch).
CC -----
CC EMBL: X94216; CAA63907.1; -
CC DR EMBL: U43142; AAB85214.1; -
CC DR EMBL: U58111; AAB02909.1; -
CC DR HSP: P15692; 1VPF.
CC DR MIM: 601528; -.
CC DR InterPro: IPR000072; -.
CC DR InterPro: IPR002400; -.
CC DR Pfam: PF00341; PDGF_1;
CC DR PRINTS: PR00438; GFCSKNOT.
CC DR PROSITE: PS00249; PDGF_1;
CC DR PROSITE: PS00278; PDGF_2; 1.
CC DR Mitogen; Growth factor; Glycoprotein; Signal; Repeat.
CC FT SIGNAL 1 ? ? POTENTIAL.
CC FT PROPEP 1 ? 102 POTENTIAL.
CC FT CHAIN 103 419 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
CC FT DOMAIN 275 365 4 X 24 AA TANDEM REPEATS.
CC FT REPEAT 275 298 1.
CC FT REPEAT 299 322 2.
CC FT REPEAT 323 346 3.
CC FT REPEAT 347 365 4 (PARTIAL).
CC FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SQ SEQUENCE 419 AA; 46883 MW; 9F598719DB3E01AF CRC64;

Query Match 100.0%; Score 539; DB 1; Length 419;
Best Local Similarity 100.0%; Pred. No. 5,1e-49;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAFESGLDLSDAEPDAGEATAYASKDLSEQLRSVSSVDLMTVLYPEYKMYKCOLRRKG 60
DB 30 AAFESGLDLSDAEPDAGEATAYASKDLSEQLRSVSSVDLMTVLYPEYKMYKCOLRRKG 89
QY 61 WOHNRQANLNSRTETIKFAAAHNTETIKSIDNEMKRTQC 102
DB 90 WOHNRQANLNSRTETIKFAAAHNTETIKSIDNEMKRTQC 131

RESULT 2
VEGC_MOUSE
ID VEGC_MOUSE STANDARD: PRT: 415 AA.
AC P97953;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR (VEGF-C) (FL14 LIGAND)
DE (FL14-L).
GN VEGFC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=97388482; PubMed=9247316;
RA Fitz L.J., Morris J.C., Towler P., Long A., Burgess P., Greco R.,
RA Wang J., Gassaway R., Nickbarg E., Kovacic S., Ciarelletta A.,
RA Giannotti J., Finerty H., Zollner R., Beier D.R., Leek L.V.,
RA Turner K.J., Wood C.R.;
RA "Characterization of murine Flt4 ligand/VEGF-C.";
```

```

RL Oncogene 15:613-618(1997).
CC -1- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
CC CELL GROWTH.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.lsb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
CC EMBL: U73620; AAC52984.1; -.
CC DR EMBL: U58112; AAB46707.1; -.
CC DR HSP: P15692; 1VPF.
CC DR MGD: MGI:109124; Vegfc.
CC DR InterPro: IPR000072; -.
CC DR InterPro: IPR002400; -.
CC DR Pfam: PF00341; PDGF_1;
CC DR PRINTS: PR00438; GFCSKNOT.
CC DR PROSITE: PS00249; PDGF_1; 1.
CC DR PROSITE: PS00278; PDGF_2; 1.
CC DR Mitogen; Growth factor; Glycoprotein; Signal; Repeat.
CC FT SIGNAL 1 ? ? POTENTIAL.
CC FT PROPEP 1 ? 98 POTENTIAL.
CC FT CHAIN 99 415 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
CC FT DOMAIN 271 361 4 X 24 AA TANDEM REPEATS.
CC FT REPEAT 271 294 1.
CC FT REPEAT 295 318 2.
CC FT REPEAT 319 342 3.
CC FT REPEAT 343 361 4 (PARTIAL).
CC FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 226 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SQ SEQUENCE 415 AA; 46471 MW; D9D3DD3CECC659D6 CRC64;

Query Match 78.8%; Score 425; DB 1; Length 415;
Best Local Similarity 78.4%; Pred. No. 4,1e-37;
Matches 80; Conservative 8; Mismatches 10; Indels 4; Gaps 1;

QY 1 AAFESGLDLSDAEPDAGEATAYASKDLSEQLRSVSSVDLMTVLYPEYKMYKCOLRRKG 60
DB 30 AAFESGLDLSDAEPDAGEATAYASKDLSEQLRSVSSVDLMTVLYPEYKMYKCOLRRKG 89
QY 61 WOHNRQANLNSRTETIKFAAAHNTETIKSIDNEMKRTQC 102
DB 90 WOHNRQANLNSRTETIKFAAAHNTETIKSIDNEMKRTQC 127

RESULT 3
GREP_HAEMIN
ID GREP_HAEMIN STANDARD: PRT: 198 AA.
AC P43732;
DT 01-NOV-1995 (Rel. 32, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GREP PROTEIN (HSP-70 COFACTOR).
GN GREP OR H10071.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fletschmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu T.-I., Giodex A., Kelley J.M.,
```

RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venier J.C.,
 RT "Whole-genome random sequencing and assembly of Haemophilus
 RT Influenzae Rd.",
 RL Science 269:496-512(1995).
 CC -I- FUNCTION: STIMULATES, JOINTLY WITH DNAK, THE ATPASE ACTIVITY OF
 CC DNAK. HELPS TO RELEASE ADP FROM DNAK THUS ALLOWING DNAK TO RECYCLE
 CC MORE EFFICIENTLY (BY SIMILARITY).
 CC -I- SIMILARITY: BELONGS TO THE GREP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U32693; AAC21750.1; ALT_INIT.
 CC DR HSPF; P09372; IDKG.
 CC DR TIGR; H10071; .
 CC DR InterPro: IPR000740; .
 CC DR Pfam: PF01025; GREP; 1.
 CC DR PRINTS: PR00773; GREP; 1.
 CC DR PROSITE: PS01071; GREP; 1.
 KW Chaperone; Heat shock.
 SQ SEQUENCE 198 AA: 22298 MW; DDCAC1B03D50EC0 CRC64;

Query Match 13.0%; Score 70; DB 1; Length 198;
 Best Local Similarity 27.7%; Pred. No. 2.2;
 Matches 23; Conservative 13; Mismatches 19; Indels 28; Gaps 4;

OY 27 LEEQLRSVSSVDELTMTVLYPEYKMYKCOLRGKGMQHNREQ-----ANLSRTTEE 76
 DB 36 LEEAIVARVLEELQ-----KTQIEEAA--NKEQDILRSRAEINLRRTREQ 81
 OY 77 TI-----KFAAHYNTIEIKSDN 95
 DB 82 DVEKAKRFALKERKSDILNTIDN 104

RESULT 4
 GC4_HUMAN STANDARD: PRT; 327 AA.
 AC P01861;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG GAMMA-4 CHAIN C REGION.
 GN IGHG4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83157104; PubMed=6299662;
 RA Ellison J.W., Buxbaum J.N., Hood L.E.;
 RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene";
 RL J. Biol. Chem. 264:11118-11124(1989).
 RN [2]
 RP SEQUENCE OF 1-30 AND 81-326.
 RX MEDLINE=70207560; PubMed=4192699;
 RA Pink J.R.L., Buttery S.H., de Vries G.M., Milstein C.;
 RT "Human immunoglobulin subclasses. Partial amino acid sequence of the
 RT constant region of a gamma 4 chain";
 RL Biochem. J. 117:33-47(1970).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: K01316; AB59394.1; ALT_INIT.
 CC DR PIR: A02150; G4HU.
 CC DR MIM: 147130; .
 CC DR InterPro: IPR000495; .
 CC DR InterPro: IPR003006; .
 CC DR Pfam: PF00047; 1g; 3.
 CC DR PROSITE: PS00290; IG_MHC; 2.
 KW Immunoglobulin domain; Immunoglobulin C region.
 FT NON_TER 1
 FT DOMAIN 1 98
 FT DOMAIN 99 110 HINGE.
 FT DOMAIN 111 220 CH2.
 FT DOMAIN 221 327 CH3.
 FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFID 27 83
 FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 141 201
 FT DISULFID 247 305
 SQ SEQUENCE 327 AA: 35940 MW; 3EDBD81EF208E7A CRC64;

Query Match 12.9%; Score 69.5; DB 1; Length 327;
 Best Local Similarity 29.4%; Pred. No. 4.5;
 Matches 25; Conservative 12; Mismatches 29; Indels 19; Gaps 3;

OY 7 LDLSDAEP-----DAGEATAYASKDLEQLRSVSSVDELTMTVLYPEY--WNKCOL 56
 DB 144 YVVSQEDPEVQPNWYVDGVEVNAKTREDFQFNSTYRVSVLTVLHODMLNGKEYKCV 203
 OY 57 RKGGMQHNREQANLSRPEETIKFA 81
 DB 204 SNKG-----LPSSIEKITSKA 219

RESULT 5
 GC51_BACSU STANDARD: PRT; 448 AA.
 AC P54376;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE PROBABLE GLYCINE DEHYDROGENASE [DECARBOXYLATING] SUBUNIT 1
 DE (EC 1.4.4.2) (GLYCINE DECARBOXYLASE) (GLYCINE CLEAVAGE SYSTEM P-
 DE PROTEIN).
 GN YOHU.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacilli; Clostridium group;
 OC Bacilli; Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX Mizuno M., Masuda S., Takemaru K., Hosono S., Sato T., Takeuchi M.,
 RA Kobayashi Y.;
 RT Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=168 / JH642;
 RA Kobayashi Y., Mizuno M., Masuda S., Takemaru K., Hosono S., Sato T.,
 RA Takeuchi M.;
 RT Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.
 CC -I- FUNCTION: THE P PROTEIN BINDS TO THE EMBL/Genbank/DBJ databases.
 CC GLYCINE. THE P PROTEIN BINDS THE ALPHA-AMINO GROUP OF GLYCINE
 CC THROUGH ITS PYRIDOXAL PHOSPHATE COFACTOR; CO(2) IS RELEASED AND
 CC THE REMAINING METHYLAMINE MOIETY IS THEN TRANSFERRED TO THE
 CC LIPONAMIDE COFACTOR OF THE H PROTEIN (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: GLYCINE + LIPIDYLPROTEIN - S-AMINOMETHYL-
 CC DIHYDROLIPOYLPROTEIN + CO(2).
 CC -1- COFACITOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- SUBUNIT: THE GLYCINE CLEAVAGE SYSTEM IS COMPOSED OF FOUR PROTEINS:
 CC P, T, L, AND H.
 CC -1- SIMILARITY: TO THE N-TERMINAL OF OTHER GLYCINE CLEAVAGE SYSTEM P-
 CC PROTEIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: D84432; BAA12547.1; -
 CC DR EMBL: Z99116; CAB14387.1; -
 CC DR Subtilast; BG11510; yphJ.
 CC KW Hypothetical protein; Oxidoreductase; Pyridoxal phosphate.
 CC SEQUENCE 448 AA: 49497 MW: 7675330F233056B8 CRC64;

Query Match 12.9%; Score 69.5; DB 1; Length 448;
 Best Local Similarity 28.4%; Pred. No. 6.6;
 Matches 23; Conservative 17; Mismatches 28; Indels 13; Gaps 4;

OY 25 KDLPEQLRS--SSVDELTMTVLYPE---YWKMYKCOLRGSGOHNEQANLSRTEETIK 79
 DB 11 KKQKMLATIGVSSIDDLFADI-PENWKYKKRKHQIKKAKSETLRETKLASKNDIVQ 69
 OY 80 FAA-----AHYNTETILKSI 93
 DB 70 YASFLGAGYDHYQPIYVDHV 90

RESULT 6
 DYHC_MOUSE STANDARD; PRT; 4644 AA.
 AC 09JHU4;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE DYNEIN HEAVY CHAIN, CYTOSOLIC (DYHC) (CYTOPLASMIC DYNEIN HEAVY CHAIN).
 GN DNMCH1 OR DNMCH1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FVB;
 RA Sasaki S., Shionoya A., Hirotsune S.;
 RT "Complete cDNA sequence of murine cytoplasmic dynein heavy chain."
 RL Submitted (JUL-2000) to the EMBL/GenBank/CDDB databases.
 CC -1- FUNCTION: DYNEIN HAS ATPASE ACTIVITY. CYTOPLASMIC DYNEIN ACTS AS A
 CC MOTOR FOR THE INTRACELLULAR RETROGRADE MOTILITY OF VESICLES AND
 CC ORGANELLES ALONG MICROTUBULES.
 CC -1- SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF
 CC INTERMEDIATE AND LIGHT CHAINS.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: AY004877; AAF91078.1; -
 CC DR MGD: MGI:103147; Dnmch1.

KW Motor protein; Microtubules; Dynein; ATP-binding; Coiled coil.
 FT DOMAIN 48 69
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 179 200
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 453 476
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 541 564
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 1169 1201
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 1229 1250
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 1355 1371
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 2012 2040
 FT MICROTUBULE-BINDING (POTENTIAL).
 FT DOMAIN 3187 3273
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 3394 3498
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 3735 3798
 FT COILED COIL (POTENTIAL).
 FT NP_BIND 1904 1911
 FT ATP (POTENTIAL).
 FT NP_BIND 2222 2229
 FT ATP (POTENTIAL).
 FT NP_BIND 2593 2600
 FT ATP (POTENTIAL).
 FT NP_BIND 2935 2942
 FT ATP (POTENTIAL).
 SQ SEQUENCE 4644 AA: 532021 MW: FE5B4E15DD479ELB CRC64;

Query Match 12.8%; Score 69; DB 1; Length 4644;
 Best Local Similarity 25.0%; Pred. No. 1.1e+02;
 Matches 27; Conservative 21; Mismatches 38; Indels 22; Gaps 5;

OY 1 AAFESGLDIS-----DAEPDAGEATAYASKDLEQLRSVSSVDELTMTVLYPEYWKMYK 53
 DB 3464 AAVEAKVNVNSTALTKLSASRERWEKT---SETFKNQNMSTIAG-DCLLSAFLIAYAGYFD 3519
 OY 54 COLRG---GMQHNREQANLSRTEETIKFAAHNTETILKSIDNEMR 98
 DB 3520 QGMROLNFTTWSHHLQQAANIQRTDIA-----RTEYLSNADERLR 3559

RESULT 7
 DYHC_RAT STANDARD; PRT; 4644 AA.
 AC P38650; Q63178;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE DYNEIN HEAVY CHAIN, CYTOSOLIC (DYHC) (CYTOPLASMIC DYNEIN HEAVY CHAIN)
 DE (MAP 1C).
 GN DNMCH1 OR DNMCH1 OR DNEC1 OR MAP1C.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WISTAR; TISSUE-Brain;
 RX MEDLINE=93376715; PubMed=7690137;
 RA Zhang Z., Tanaka Y., Nonaka S., Aizawa H., Kawasaki H., Nakata T.,
 RA Hirokawa N.;
 RT "The primary structure of rat brain (cytoplasmic) dynein heavy chain,
 RT a cytoplasmic motor enzyme."
 RL Proc. Natl. Acad. Sci. U.S.A. 90:7928-7932(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-Brain;
 RX MEDLINE=93264075; Pubmed=7684232;
 RA Miyami A., Paschal B.W., Mazumdar M., Vallée R.B.;
 RT "Molecular cloning of the retrograde transport motor cytoplasmic
 RT dynein (MAP 1C)."
 RL Neuron 10:787-796(1993).
 CC -1- FUNCTION: DYNEIN HAS ATPASE ACTIVITY. CYTOPLASMIC DYNEIN ACTS AS A
 CC MOTOR FOR THE INTRACELLULAR RETROGRADE MOTILITY OF VESICLES AND
 CC ORGANELLES ALONG MICROTUBULES.
 CC -1- SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF
 CC INTERMEDIATE AND LIGHT CHAINS.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: D13896; BAA02996.1; -
DR EMBL: L08505; AAA41103.1; -
DR PIR: A38905; A38905.

KW Motor protein; Microtubules; Dynein; ATP-binding; Coiled coil.
FT DOMAIN 48 69
FT DOMAIN 179 200 COILED COIL (POTENTIAL).
FT DOMAIN 453 476 COILED COIL (POTENTIAL).
FT DOMAIN 541 564 COILED COIL (POTENTIAL).
FT DOMAIN 1169 1201 COILED COIL (POTENTIAL).
FT DOMAIN 1229 1250 COILED COIL (POTENTIAL).
FT DOMAIN 1355 1371 COILED COIL (POTENTIAL).
FT DOMAIN 2012 2040 MICROTUBULE-BINDING (POTENTIAL).
FT DOMAIN 3187 3273 COILED COIL (POTENTIAL).
FT DOMAIN 3394 3498 COILED COIL (POTENTIAL).
FT DOMAIN 3735 3798 COILED COIL (POTENTIAL).
FT NP_BIND 1904 1911 ATP (POTENTIAL).
FT NP_BIND 2222 2229 ATP (POTENTIAL).
FT NP_BIND 2593 2600 ATP (POTENTIAL).
FT NP_BIND 2935 2942 ATP (POTENTIAL).
FT CONFLICT 1024 1025 SR -> MP (IN REF. 2).
FT CONFLICT 1772 1772 N -> D (IN REF. 2).
FT CONFLICT 2098 2098 P -> A (IN REF. 2).
FT CONFLICT 2139 2139 F -> V (IN REF. 2).
FT CONFLICT 2175 2175 D -> A (IN REF. 2).
FT CONFLICT 2185 2185 K -> Q (IN REF. 2).
FT CONFLICT 2366 2366 L -> V (IN REF. 2).
FT CONFLICT 2382 2382 T -> S (IN REF. 2).
FT CONFLICT 2463 2463 G -> A (IN REF. 2).
FT CONFLICT 3219 3219 A -> D (IN REF. 2).
FT CONFLICT 4131 4131 R -> K (IN REF. 2).
FT CONFLICT 4366 4366 F -> S (IN REF. 2).
FT CONFLICT 4511 4511 A -> G (IN REF. 2).
SQ SEQUENCE 4644 AA; 532240 MW; 8C6ABDBEDF75D82 CRC64;

Query Match 12.8%; Score 69; DB 1; Length 4644;
Best Local Similarity 25.0%; Pred. No. 1.1e+02;
Matches 27; Conservative 21; Mismatches 38; Indels 22; Gaps 5;

QY 1 AAFESGLDLS-----DDEPDAGEATATASKDLBEOQLSSVDELMVLYPEIKWKMK 53
DB 3464 AAVEAVNKSSTALSLKSLAERERWERT---SEPFKNQMSITAG-DCLLSAEFAIYAGYRD 3519
QY 54 COLRKG---GMOHNEQOANLNSRTETKEFAAHVNTLEIKSIDNEMR 98
DB 3520 QOMRONLFTTWSHHCQANIOFRTDIA-----RTEYLSNADERLR 3559

RESULT 8
ANX7_XENLA
ID ANX7_XENLA STANDARD; PRT; 512 AA.
AC 092125;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ANNEXIN A7 (ANNEXIN VII) (SYNEXIN).
GN ANNEXIN A7 OR ANX7.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenoportidae; Xenopus.
OX NCBI_TaxID=8335;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96265035; Pubmed=8670145;
RX Srivastava M., Zhang-Kock Z.Y., Caohuy H., McPhie P., Pollard H.B.;

RT "Novel isoforms of synexin in Xenopus laevis: multiple tandem PGOM
RT repeats distinguish mRNAs in specific adult tissues and embryonic
RT stages.";
RL Biochem. J. 316:729-735(1996).
CC -1- FUNCTION: CALCIUM/PHOSPHOLIPID-BINDING PROTEIN WHICH PROMOTES
CC MEMBRANE FUSION AND IS INVOLVED IN EXOCYTOSIS.
CC -1- SIMILARITY: BELONGS TO THE ANNEXIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: U16365; AAB18145.1; -
DR HSSP: P26256; LAEI.
DR InterPro: IPR001464; -
DR Pfam: PF00191; annexin; 4.
DR PRINTS: PR00196; ANNEXIN.
DR PROSITE: PS00223; ANNEXIN; 3.
KW Annexin; Calcium/phospholipid-binding; Repeat.
FT REPEAT 218 277 ANNEXIN 1.
FT REPEAT 290 350 ANNEXIN 2.
FT REPEAT 373 433 ANNEXIN 3.
FT REPEAT 449 509 ANNEXIN 4.
SQ SEQUENCE 512 AA; 53314 MW; 70D532D2524388A4 CRC64;

Query Match 12.7%; Score 68.5; DB 1; Length 512;
Best Local Similarity 31.4%; Pred. No. 9.8;
Matches 27; Conservative 13; Mismatches 37; Indels 9; Gaps 4;

QY 20 TAYASKDLBEOQLSSV--VDELMVLY-----YPEYKWKCOLRKGWQHNEQOANLNS 72
DB 252 TAY-GMDLKKDLKSLSSVVEELITAFSPSTYYDMSLYNA-MGAGTQGEVLEILCT 309
QY 73 RTEETIKFAAHVNTLEIKSIDNEMR 98
DB 310 RNSLEIRNIVACYKQEFGEIEKDIR 335

RESULT 9
RIR2_TOBAC
ID RIR2_TOBAC STANDARD; PRT; 329 AA.
AC P49730;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE SMALL CHAIN (EC 1.17.4.1)
DE (RIBONUCLEOTIDE REDUCTASE) (R2 SUBUNIT).
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; eusterids I;
OC Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. BRIGHT YELLOW 2;
RA Phillips G., Chaboute M.E., Clement B., Gigot C.;
RL Submitted (OCT-1995) to the EMBL/Genbank/DOBJ databases.
CC -1- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: 2'DEOXYRIBONUCLEOSIDE DIPHOSPHATE + OXIDIZED
CC THIOREDOXIN + H(2)O = RIBONUCLEOSIDE DIPHOSPHATE + REDUCED
CC THIOREDOXIN.
CC -1- COFACTOR: CONTAINS TWO IRON IONS (BY SIMILARITY).
CC -1- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.
CC -1- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL CHAIN.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
CC SMALL CHAIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X92443; CAA63194.1; -
DR HSSP: P11157; 1XSM.
DR InterPro: IPR000358; -
DR Pfam: PF00268; ribonuc_red. 1.
DR PROSITE: PS00368; RIBORED_SMALL; 1.
KW Oxidoreductase; DNA replication; Iron.
FT METAL 75 75 IRON 1 (BY SIMILARITY).
FT METAL 106 106 IRON 1 AND 2 (BY SIMILARITY).
FT METAL 109 109 IRON 1 (BY SIMILARITY).
FT METAL 168 168 IRON 2 (BY SIMILARITY).
FT METAL 202 202 IRON 2 (BY SIMILARITY).
FT METAL 205 205 IRON 2 (BY SIMILARITY).
FT ACT_SITE 113 113 BY SIMILARITY.
SQ SEQUENCE 329 AA; 37608 MW; 65E7AA183AF045FD CRC64;

Query Match 12.4%; Score 67; DB 1; Length 329;
Best Local Similarity 34.3%; Pred. No. 8.3;
Matches 23; Conservative 11; Mismatches 25; Indels 8; Gaps 5;

OY 28 EQLSVSSVDE--LMTVLYPEYKMKYKCOLRKGQHNREQANLSRPE--ETTFEAAA 83
DB 6 EEPFLA-SSPDRFCMPPIQYPOIWEYKKAL-ASEW--TAEEVDLSSDRHMETLTPGER 61
OY 84 HNTTEL 90
DB 62 HETTHVL 68

RESULT 10

Y137_CHLPN STANDARD; PRT; 251 AA.
AC Q92946; 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN CPN0137/CP0635.
GN CPN0137 OR CP0635.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CML029.
RX MEDLINE=99206606; PubMed=10192388;
RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.,
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Umayam L.A., Utterback T.,
RA Berry K., Bass S., Linher K., Weidman J., Khouri H., Craven B.,
RA Bowman C., Dooson R., Winn M., Nelson W., Debo R., Kolonay J.,
RA McClarty G., Salberg S.L., Eisen J., Fraser C.M.,
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;

RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kohara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CML029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -1- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AE001600; AAD18290.1; -
DR EMBL: AE002221; AAF38450.1; -
DR EMBL: AP002545; BAA98347.1; -
DR TIGR: CP0635; -
DR InterPro: IPR002678; -
DR Pfam: PF01784; DUF34.1.
KW Hypothetical protein
SQ SEQUENCE 251 AA; 27236 MW; 6D609CDE3AAB13EF CRC64;

Query Match 12.2%; Score 66; DB 1; Length 251;
Best Local Similarity 22.3%; Pred. No. 7.7;
Matches 21; Conservative 19; Mismatches 48; Indels 6; Gaps 2;

OY 5 SGLDSDAEPDAGEATVAVASKDLERSSVDELMTVLYPEYKMKYKCOLRKGQHN 64
DB 25 NGLQVGDPTQPYKIAVAVATDLETKQVAEAVNLYIHGIFPKMGMPYPT--GMHK 82
OY 65 REQANINSTRTEETIKFAAHYNTTELKSIDNEWR 98
DB 83 RIQLDI---EHNQIDLVHPLDAPHTLGNMR 112

RESULT 11

CPSM_HUMAN STANDARD; PRT; 1500 AA.
AC P31327; 043774;
DT 01-JUL-1993 (Rel. 26, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CARBAMOYL-PHOSPHATE SYNTHASE [AMMONIA], MITOCHONDRIAL PRECURSOR
DE (EC 6.3.4.16) (CARBAMOYL-PHOSPHATE SYNTHETASE I) (CPSASE I).
GN CPSI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92084128; PubMed=1840546;
RA Haraoguchi Y., Uchino T., Takiguchi M., Endo F., Mori M.,
RA Matsuda I.,
RT "Cloning and sequence of a cDNA encoding human carbamyl phosphate
RT synthetase I: molecular analysis of hyperammonemia.";
RL Gene 107:335-340(1991).
RN [2]
RP SEQUENCE FROM N.A., VARIANT CPSI DEF. MET-544, AND VARIANT ALA-344.

RC TISSUE=Liver;
RX MEDLINE=98375696; PubMed=9711878;
RA Finch U., Kohlschuetter A., Schaefer H., Sperhake K., Colombo J.-P.,
RA Gal A.,
RT "Prenatal diagnosis of carbamyl phosphate synthetase I deficiency by
RT identification of a missense mutation in CPSI.";
RL Hum. Mutat. 12:206-211(1998).
CC -1- FUNCTION: INVOLVED IN THE UREA CYCLE OF UREOTELIC ANIMALS WHERE
CC THE ENZYME PLAYS AN IMPORTANT ROLE IN REMOVING EXCESS AMMONIA
CC FROM THE CELL.

CC -1- CATALYTIC ACTIVITY: 2 ATP + NH(3) + CO(2) + H(2)O = 2 ADP +
 CC ORTHOPHOSPHATE + CARBAMOYL PHOSPHATE.
 CC -1- ENZYME REGULATION: REQUIRES N-ACETYLGLUTAMATE AS AN ALLOSTERIC
 CC ACTIVATOR.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.
 CC -1- TISSUE SPECIFICITY: PRIMARILY IN THE LIVER AND SMALL INTESTINE.
 CC -1- DISEASE: DEFECTS IN CPS1 ARE THE CAUSE OF AN AUTOSOMAL RECESSIVE
 CC METABOLIC DISORDER THAT CAUSE A TYPE OF HYPERAMMONEMIA. CLINICAL
 CC SYMPTOMS ARE VOMITING IN INFANCY, PROTEIN INTOLERANCE,
 CC INTERMITTENT ATAXIA, SEIZURES, LETHARGY, AND MENTAL RETARDATION.
 CC -1- SIMILARITY: TO OTHER CARBAMOYL-PHOSPHATE SYNTHETASES. ALSO
 CC CONTAINS A GLUTAMINE AMIDOTRANSFERASE-LIKE DOMAIN BUT THE Cysteine
 CC RESIDUE WHICH IS ESSENTIAL FOR AMIDO TRANSFERASE ACTIVITY, IS
 CC MISSING IN THIS PROTEIN SO THIS DOMAIN IS NO LONGER ACTIVE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: D90282; BA014328.1; -;
 CC EMBL: Y15793; CA05785.1; -;
 CC PIR: J01348; J01348.
 CC HSSP: P00968; LUDB.
 CC MEROPS: M38.971; -;
 CC MIM: 237300; -;
 CC InterPro: IPR000901; -;
 CC InterPro: IPR000991; -;
 CC InterPro: IPR001317; -;
 CC InterPro: IPR002474; -;
 CC Pfam: PF00289; CPSase_L_chain; 2.
 CC Pfam: PF00986; CPSase_sm_chain; 1.
 CC Pfam: PF00117; GATase; 1.
 CC PRINTS: PR00098; CPSASE.
 CC PRINTS: PR00099; CPSGATASE.
 CC PROSITE: PS00866; CPSASE_1; 2.
 CC PROSITE: PS00867; CPSASE_2; 2.
 CC LIGASE; Duplication; Transist peptide; Mitochondrion; ATP-binding;
 CC Urea cycle; Polymorphism; Disease mutation.
 CC TRANSIT 38
 CC CHAIN 1 38
 CC NP_BIND 571 1500
 CC NP_BIND 718 768
 CC NP_BIND 1113 1171
 CC NP_BIND 1259 1302
 CC DOMAIN 39 219
 CC DOMAIN 220 410
 CC DOMAIN 411 1500
 CC REPEAT 419 876
 CC REPEAT 970 1410
 CC VARIANT 344 344
 CC VARIANT 544 544
 CC CONFLICT 111 111
 CC CONFLICT 279 279
 CC CONFLICT 338 338
 CC CONFLICT 718 722
 CC CONFLICT 729 729
 CC CONFLICT 749 749
 CC CONFLICT 1161 1162
 CC CONFLICT 1204 1205
 CC CONFLICT 1254 1254
 CC CONFLICT 1266 1266
 CC CONFLICT 1283 1283
 CC CONFLICT 1303 1303
 CC CONFLICT 1406 1406
 CC SEQUENCE 1500 AA; 164939 MW; E53A2D77563961D CRC64;

Query Match 12.2%; Score 66; DB 1; Length 1500;
 Best Local Similarity 28.2%; Pred. No. 63;
 Matches 24; Conservative 13; Mismatches 28; Indels 20; Gaps 4;
 QY 5 SCGLSDAPDGDGEATVA-SKOLEOLRSVSVELMTVLPEYWKMKC-----OL 56
 DB 835 SNLDRLKELSEPSSTRIVAIKAIIDNM-----SLDEIKETIYIDKFLYKMDILMENT 890
 QY 57 RRGGMQHNREQANLNSRPEETIKFA 81
 DB 891 LKG-----LNSESMPEETLKRA 907
 RESULT 12
 ID SYEA_YEAST STANDARD; PRT; 502 AA.
 AC P15625;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PHENYLALANYL-TRNA SYNTHETASE ALPHA CHAIN (EC 6.1.1.20) (PHENYLALANINE-
 DE -TRNA LIGASE ALPHA CHAIN) (PHERS).
 GN FR52 OR YFL022C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=89008440; Pubmed=3049607;
 RA Sanni A., Mirande M., Ebel J.-P., Boulanger Y., Waller J.-P.,
 RA Pascolo F.,
 RT Structure and expression of the genes encoding the alpha and beta
 RT subunits of yeast phenylalanyl-trna synthetase.";
 RL J. Biol. Chem. 263:15407-15415(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RX MEDLINE=95400292; Pubmed=7670463;
 RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
 RA Sasasuna S.-I., Sasasuna M., Tsuchiya Y., Soeda E., Yokoyama K.,
 RA Yamazaki M., Tashiro H., Eki T.;
 RT "Analysis of the nucleotide sequence of chromosome VI from
 RT Saccharomyces cerevisiae.";
 RL Nat. Genet. 10:261-268(1995).
 RN [3]
 RP PARTIAL SEQUENCE, AND TRNA-BINDING DOMAIN.
 RX MEDLINE=89121098; Pubmed=2644133;
 RA Pascolo F., Sanni A., Potier S., Ebel J.-P., Boulanger Y.,
 RT "Identification of the major tRNA(Phe) binding domain in the
 RT tetrameric structure of cytoplasmic phenylalanyl-trna synthetase from
 RT baker's yeast.";
 RL FEBS Lett. 242:351-356(1989).
 RN [4]
 RP ACETYLATION.
 RA Gattarel J.I., Fletcher B., Kobayashi R., Latter G.I., Schwender B.,
 RA Volpe T., Warner J.R., McLaughlin C.S.;
 RL Submitted (SEP-1994) to the SWISS-PROT data bank.
 CC -1- CATALYTIC ACTIVITY: ATP + L-PHENYLALANINE + TRNA(PHE) = AMP +
 CC PYROPHOSPHATE + L-PHENYLALANYL-TRNA(PHE).
 CC -1- SUBUNIT: Tetramer of two alpha and two beta chains.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC PHE-TRNA SYNTHETASE ALPHA CHAIN SUBFAMILY 2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```

CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: J03965; AAA5152.1; -.
CC EMBL: D50617; BAA09216.1; -.
CC PIR: B31990; YEBYAC.
CC HSSP: P27001; 1PYS.
CC YEPD: 7545; -.
CC SGD: S0001872; FR52.
CC InterPro: IPR002106; -.
CC InterPro: IPR002319; -.
CC Pfam: PF01409; tRNA-synt_2d.1.
CC PROSITE: PS00179; AA_TRNA_LIGASE_II.1.
CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Acetylation.
CC INIT_MET 0 0
CC MOD_RES 1 1 ACETYLATION.
CC DOMAIN 1 172 CONSTRAINTS THE MAJOR TRNA-PHE BINDING
CC CONFLICT 176 176 N -> S (IN REF. 1).
CC CONFLICT 288 288 D -> E (IN REF. 1).
CC SEQUENCE 502 AA; 57380 MW; BDB7EAC675AB4037 CRC64;

Query Match
Best Local Similarity 12.1%; Score 65; DB 1; Length 502;
Matches 28; Conservative 19; Mismatches 46; Indels 26; Gaps 4;

OY 7 LDLSAEPDAGATVAVASDLE---QLRSVSSVDEL-----MTVLYPERKMKYKCQ 55
DB 45 LEFSVDTVTYDLTREKQOILNESSYEIKLVLIQELGOLIKIDVMSKIPQGVKGVAR 104
OY 56 LKRGGM-----OHNRQANILNRTETIKFAAAHYNTILKSID--NEWRK 99
DB 105 AFKNQWIKKNAENLELSAKLQNTDLNLTDTQSLAOKKNSHLSDIAKLINDLKR 163

RESULT 13
PROT YEAST STANDARD: PRT; 2413 AA.
AC P33334;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE PRE-MRNA SPLICING FACTOR PRP8.
DE PRP8 OR RNA8 OR DBF3 OR DNA39 OR YHR165C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95304847; PubMed=7785334;
RA Hodges P.E., Jackson S.P., Brown J.D., Beggs J.D.;
RT "Extraordinary sequence conservation of the PRP8 splicing factor.";
RL Yeast 11:337-342(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95140615; PubMed=7838707;
RA Shea J.E., Toyon J.H., Johnston L.H.;
RT "The budding yeast U5 snRNP PRP8 is a highly conserved protein which
RT links RNA splicing with cell cycle progression.";
RL Nucleic Acids Res. 22:5555-5564(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
RA Nhan M., Rifkin L., Riles L., St Peter H., Trevisan E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;

```

```

RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII.";
RT Science 265:2077-2082(1994).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=88216580; PubMed=2835658;
RA Jackson S.P., Lossky M., Beggs J.D.;
RT "Cloning of the RNA8 gene of Saccharomyces cerevisiae, detection of
RT the RNA8 protein, and demonstration that it is essential for nuclear
RT pre-mRNA splicing.";
RT Mol. Cell. Biol. 8:1067-1075(1988).
CC -1- FUNCTION: INVOLVED IN PRE-MRNA SPLICING. U5 snRNP PROTEIN.
CC APPEARS TO CONTACT THE PRE-MRNA DURING SPLICING. ALSO HAS A ROLE
CC IN CELL CYCLE. BINDS RNA.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: STRONG, TO C.ELEGANS PROTEIN C50C3.6.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: 224732; CAA80854.1; -.
CC EMBL: L29421; AAB67044.1; -.
CC EMBL: U00027; AAB68011.1; -.
CC PIR: S34670; S34670.
CC DR PIR: S48905; S48905.
CC SGD: S0001208; PRP8.
CC mRNA processing; mRNA splicing; Spliceosome; Nuclear protein;
CC RNA-binding.
CC DOMAIN 5 9 POLY-PRO.
CC FT DOMAIN 20 27 POLY-PRO.
CC FT DOMAIN 50 56 POLY-PRO.
CC FT DOMAIN 72 78 POLY-PRO.
CC FT CONFLICT 388 420 PHLYNSRPRSVRIPIWNPVSCIONDEEYDTP ->
CC LIIIPGVQCAVYGIILIQCVLSRTMSTRL
CC FT CONFLICT 1132 1132 T -> S (IN REF. 2).
CC FT CONFLICT 1575 1575 W -> C (IN REF. 2).
CC SO SEQUENCE 2413 AA; 279501 MW; 8FAF6F89D34D3508 CRC64;

Query Match
Best Local Similarity 12.1%; Score 65; DB 1; Length 2413;
Matches 17; Conservative 13; Mismatches 20; Indels 10; Gaps 2;

OY 50 KMYKCOLRKGWQHNRQEA-----NLNSTRTEETIKFAAAHYNTILKSIDNEWKRT 100
DB 2097 KNAELEAARSEKQNDENEAAGASTVKTITNAGGEIIVVASADYESQTFSS-KNEWKRS 2155

RESULT 14
FOS_CYPCA
ID FOS_CYPCA STANDARD: PRT; 347 AA.
AC P79702;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE P55-C-FOS PROTO-ONCOGENE PROTEIN (CELLULAR ONCOGENE C-FOS).
GN FOS.
OS Cypripus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprininae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RA Chang M.S., Huang C.J.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: NUCLEAR PHOSPHOPROTEIN WHICH FORMS A TIGHT BUT NON-

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:50:49 ; Search time 128.06 Seconds
(without alignments)
105.361 Million cell updates/sec

Title: US-09-427-657-2_COPY_30_131
Perfect score: 539
Sequence: 1 AAFESGDLSDPEPDAGEAR.....AHNTTELKSIDNEMKRTQC 102

Scoring table:
BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPTREMBL_16:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mmc:*
9: sp.phage:*
10: sp.plant:*
11: sp.potent:*
12: sp.unclassified:*
13: sp.vertebrate:*
14: sp.virus:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	468	86.8	420	6 Q9XS50	Q9XS50 bos taurus
2	421	78.1	418	13 Q57352	Q57352 coturnix co
3	165	30.6	326	11 Q35251	Q35251 rattus norv
4	163	30.2	358	11 P97946	P97946 mus musculu
5	155	28.8	354	4 Q43915	Q43915 homo sapien
6	77	14.3	1494	11 Q88902	Q88902 rattus norv
7	74	13.7	1345	4 Q9H357	Q9H357 homo sapien
8	74	13.7	1636	4 Q9H357	Q9H357 homo sapien
9	70.5	13.1	146	11 Q60482	Q60482 cavia porce
10	70.5	13.1	1729	3 Q10287	Q10287 schizosacch
11	70.5	13.1	2228	10 Q48579	Q48579 arabidopsi
12	70.5	13.1	4488	11 Q9QZHI	Q9QZHI mus musculu
13	70	13.0	371	2 Q56080	Q56080 salmonella
14	69	13.0	371	2 Q56080	Q56080 salmonella
15	69	12.8	124	2 Q05206	Q05206 streptococ
16	69	12.8	786	5 Q21027	Q21027 caenorhabd
17	69	12.8	1403	10 Q9SUZ1	Q9SUZ1 arabidopsi
18	69	12.8	2087	4 Q9Y4G5	Q9Y4G5 homo sapien
19	68.5	12.7	984	2 Q9PMEL	Q9PMEL campylobact

20	68.5	12.7	1265	11 Q61992	Q61992 mus musculu
21	68.5	12.7	3053	11 Q9ER9	Q9ER9 mus musculu
22	68	12.6	489	5 Q9A28	Q9A28 caenorhabd
23	68	12.6	611	2 Q45855	Q45855 coxiella bu
24	68	12.6	612	2 Q45932	Q45932 coxiella bu
25	68	12.6	706	2 Q52870	Q52870 coxiella bu
26	67.5	12.5	347	5 Q9VMY5	Q9VMY5 drosophila
27	67	12.4	77	2 Q54882	Q54882 s putative
28	67	12.4	234	14 Q55722	Q55722 chilo iride
29	67	12.4	375	10 Q23566	Q23566 arabidopsi
30	67	12.4	453	2 Q9UZV9	Q9UZV9 neisseria m
31	67	12.4	453	2 Q9JV08	Q9JV08 neisseria m
32	67	12.4	493	5 Q01798	Q01798 caenorhabd
33	67	12.4	532	2 Q9RYV1	Q9RYV1 delnoccoc
34	67	12.4	703	2 Q9KQZ2	Q9KQZ2 vibrio chol
35	67	12.4	1299	10 Q9STE0	Q9STE0 brassica na
36	66.5	12.3	602	5 Q9W397	Q9W397 drosophila
37	66.5	12.3	689	2 Q9FCS2	Q9FCS2 neisseria m
38	66	12.2	302	10 Q9FLP3	Q9FLP3 arabidopsi
39	66	12.2	441	4 Q9H993	Q9H993 homo sapien
40	65.5	12.2	532	10 Q80941	Q80941 arabidopsi
41	65	12.1	239	2 Q9RV15	Q9RV15 delnoccoc
42	65	12.1	246	4 Q9UFR5	Q9UFR5 homo sapien
43	65	12.1	1612	5 Q9VYQ2	Q9VYQ2 drosophila
44	65	12.1	2408	5 Q22184	Q22184 caenorhabd
45	64.5	12.0	148	4 Q75575	Q75575 homo sapien

ALIGNMENTS

RESULT	ID	Q9XS50	PRELIMINARY:	PRT:	420 AA.
Q9XS50	AC	Q9XS50			
DT	01-NOV-1999	(TRENBLrel. 12, Created)			
DT	01-NOV-1999	(TRENBLrel. 12, Last sequence update)			
DT	01-MAR-2001	(TRENBLrel. 16, Last annotation update)			
DE	VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR.				
OS	Bos taurus (Bovine).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OC	Bovidae; Bovinae; Bos.				
NC	NCBI_TaxID=9913;				
RM	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=HEART;				
RA	Liu X., Yonekura H., Yamagishi S., Yamamoto Y., Yamamoto H.;				
RT	"Structure and expression of bovine VEGF family.";				
RT	Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.				
DR	EMBL; AB004275; BAA77687.1; -				
DR	HSSP; P13692; IVP.				
DR	InterPro; IPR000072; -				
DR	Pfam; PF00341; PDGF_1.				
DR	PROSITE; PS00249; PDGF_1; 1.				
DR	PROSITE; PS0278; PDGF_2; 1.				
DR	SMART; SM00141; PDGF_1.				
KM	Signal.				
FT	SIGNAL	1	20	POTENTIAL.	
FT	CHAIN	21	420	VASCULAR ENDOTHELIAL GROWTH FACTOR C.	
FT	SEQUENCE	420 AA;	46681 MM;	58BA84317A3C8E2D C8C64;	
Query Match	Best Local Similarity	86.8%;	Score 468;	DB 6;	Length 420;
Matches	Conservative	85.3%;	Pred. No. 8.4e-41;		
	Matches	87;	Conservative	5;	Mismatches 10;
				Indels	0;
				Gaps	0;
Qy	1	AAFEGLSDPEPDAGATVASKDLEQLRSVSVDELMTVLYPEYWKWKCOLRRGG	60		
Db	31	AAFEGLFSDPEPDAGENKAYAGKEMEQLRSSVDELMTVLYPEYWKWKCOLRRGG	90		
Qy	61	WQHNREQANLSRTEETIKFAAHNTTELKSIDNEMKRTQC	102		

DB 91 WOHSTEQANTNIRTEGTEKFAAAHYNTTEILRSIDNEMRKTQC 132

RESULT 2

ID 057352 PRELIMINARY: PRT: 418 AA.

AC 057352;

DT 01-JUN-1998 (TREMBLrel. 06, Created)

DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR.

GN VEGF-C.

OS Coturnix coturnix japonica (Japanese quail).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Coturnix.

NCBI_TaxID=93934;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98167900; PubMed=9435294;

RA Eichmann A., Corbel C., Jaffredo T., Breant V., Joukov V., Kumar V.,

RA Allitalo K., Le Douarin N.M.;

RT "Avian VEGF-C: cloning, embryonic expression pattern and stimulation

RT of the differentiation of VEGFR2-expressing endothelial cell

RT precursors."

RL Development 125:743-752(1998).

DR EMBL: Y15837; CAA75799.1; -.

DR HSSP: P15692; IVP.

DR InterPro: IPR000072; -.

DR InterPro: IPR002400; -.

DR Pfam: PF00341; PDGF; 1.

DR PRINTS: PR00438; GFCYSKNOT.

DR ProDom: PD001629; -; 1.

DR PROSITE: PS00249; PDGF_1; 1.

DR PROSITE: PS50278; PDGF_2; 1.

DR SMART: SM00141; PDGF; 1.

KW Signal.

FT SIGNAL 1 31 POTENTIAL.

FT CHAIN 111 418 VASCULAR ENDOTHELIAL GROWTH FACTOR C.

SEQ SEQUENCE 418 AA; 46839 MW; 099BFC9151BFB2 CRC64;

Query Match 78.1%; Score 421; DB 13; Length 418;

Best Local Similarity 75.5%; Pred. No. 6.6e-36;

Matches 77; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

OY 1 AAFESGLDSDADPGAGTAYASKDLFEOLRSVSSVDLMTVLVPEYKMYKCOLRKG 60

DB 29 AAYESGHGTYEEEPGAGPKAHASKDLFEOLRSVSSVDLMTVLVPEYKMYKCOLRKG 88

OY 61 WOHNRBOANLNSRTEETIKFAAAHYNTTEILRSIDNEMRKTQC 102

DB 89 WOHNRHSSSDJRSDSLKFAAAHYNAELTKSIDNEMRKTQC 130

RESULT 3

ID 035251 PRELIMINARY: PRT: 326 AA.

AC 035251;

DT 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE VASCULAR ENDOTHELIAL GROWTH FACTOR D.

GN VEGF-D.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE DAWLEY;

RX MEDLINE=97349118; PubMed=9205122;

RX Yamada Y., Nezu J., Shimane M., Hirata Y.;

RT "Molecular cloning of a novel vascular endothelial growth factor,

RT VEGF-D."

RL Genomics 42:483-488(1997).

DR EMBL: AF014827; AAB6557.1; -.

DR HSSP: P15692; IVP.

DR InterPro: IPR000072; -.

DR Pfam: PF00341; PDGF; 1.

DR ProDom: PD001629; -; 1.

DR PROSITE: PS00249; PDGF_1; 1.

DR PROSITE: PS50278; PDGF_2; 1.

DR SMART: SM00141; PDGF; 1.

SEQ SEQUENCE 326 AA; 37112 MW; 1261AFA373596C00 CRC64;

Query Match 30.6%; Score 165; DB 11; Length 326;

Best Local Similarity 40.3%; Pred. No. 2.4e-09;

Matches 31; Conservative 22; Mismatches 16; Indels 8; Gaps 2;

OY 28 EOLRSVSSVDLMTVLVPEYKMYKCOLRKGWOHNRBOANLNSR--TEETIKFAAAHY 85

DB 46 EQIRASTLEELLOVHASEDKMLMRCRLK-----KSLAVDSRSTSHRSTREAFATFY 99

OY 86 NTEILKSIDNEMRKTQC 102

DB 100 DTEILKVIDEWMORTQC 116

RESULT 4

ID P97946 PRELIMINARY: PRT: 358 AA.

AC P97946;

DT 01-MAY-1997 (TREMBLrel. 03, Created)

DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE VASCULAR ENDOTHELIAL GROWTH FACTOR D (C-FOS INDUCED GROWTH FACTOR).

GN VEGF-D OR FlGF.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J;

RX MEDLINE=97030254; PubMed=8876195;

RA Ordanini M., Marconini L., Ferruzzi R., Oliviero S.;

RT "Identification of a c-fos-induced gene that is related to the

RT platelet-derived growth factor/vascular endothelial growth factor

RT family."

RL Proc. Natl. Acad. Sci. U.S.A. 93:11675-11675(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=LUNG;

RX MEDLINE=97349118; PubMed=9205122;

RA Yamada Y., Nezu J., Shimane M., Hirata Y.;

RT "Molecular cloning of a novel vascular endothelial growth factor,

RT VEGF-D."

RL Genomics 42:483-488(1997).

DR EMBL: X99572; CAA67892.1; -.

DR EMBL: D89628; BAA14002.1; -.

DR HSSP: P15692; IVP.

DR MGD: MGI:108037; FlGF.

DR InterPro: IPR000072; -.

DR Pfam: PF00341; PDGF; 1.

DR ProDom: PD001629; -; 1.

DR PROSITE: PS00249; PDGF_1; 1.

DR PROSITE: PS50278; PDGF_2; 1.

DR SMART: SM00141; PDGF; 1.

SEQ SEQUENCE 358 AA; 40908 MW; 6636B17FDF07037C CRC64;

Query Match 30.2%; Score 163; DB 11; Length 358;

Best Local Similarity 39.0%; Pred. No. 4.3e-09;

Matches 30; Conservative 23; Mismatches 16; Indels 8; Gaps 2;

RA Nagase T., Kikuno R., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human

RESULT	10		
Q10287		PRELIMINARY;	PTT; 1729 AA.
ID	Q10287		
AC	Q10287;		
DT	01-NOV-1996 (TrEMBLrel. 01, Created)		
DT	01-NOV-1996 (TrEMBLrel. 01, last sequence update)		
DT	01-MAR-2001 (TrEMBLrel. 16, last annotation update)		
DE	1.3-BETA-GLUCAN SYNTHASE COMPONENT CPS1 (EC 2.4.1.34) (1.3-BETA-D-GLUCAN-UDP GLUCOSYLTRANSFERASE).		
GN	CPS1 OR DRC1 OR SPEC1G7.05C.		
OC	Schizosaccharomyces pombe (fission yeast).		
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;		
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;		
OX	NCBI_TaxID=4896;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	SPRAIN-972;		
RX	MEDLINE=98062194; PubMed=9401022;		
RA	Isaliguro J., Salton A., Duran A., Ribas J.C.;		
RT	"cps1", a Schizosaccharomyces pombe gene homolog of Saccharomyces cerevisiae FKS genes whose mutation confers hypersensitivity to cyclosporin A and papulacandin B.";		
RT	J. Bacteriol. 179:7653-7662(1997).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	SPRAIN-972;		
RA	Lyne M., Rajandream M.A., Barrell B.G., Devlin K., Churcher C.M.;		
RT	Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.		
CC	-1- CATALYTIC ACTIVITY: UDP-GLUCOSE + (1.3-BETA-D-GLUCOSYL)(N) = UDP + (1.3-BETA-D-GLUCOSYL)(N+1).		
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).		
CC	EMBL: D78352; BA011369.1; -.		
DR	EMBL: AL021839; CAA17059.1; -.		
DR	InterPro: IPR003440; -.		
DR	Pfam: PF02364; Glucan_synthase; 1.		
KW	Transmembrane; Transferase; Glycosyltransferase.		
FT	TRANSMEM 378 398		POTENTIAL.
FT	TRANSMEM 417 437		POTENTIAL.
FT	TRANSMEM 448 468		POTENTIAL.
FT	TRANSMEM 1180 1200		POTENTIAL.
FT	TRANSMEM 1237 1257		POTENTIAL.
FT	TRANSMEM 1359 1379		POTENTIAL.
FT	TRANSMEM 1440 1460		POTENTIAL.

FT TRANSMEM 1484 1504 POTENTIAL.
 FT DOMAIN 1495 1498 POLY-PHE.
 FT TRANSMEM 1678 1698 POTENTIAL.
 SQ SEQUENCE 1729 AA; 199668 MW; B3CABD2960201AC1 CRC64;

Query Match 13.1%; Score 70.5; DB 3; Length 1729;
 Best Local Similarity 27.4%; Pred. No. 1.2e+02;
 Matches 29; Conservative 14; Mismatches 42; Indels 21; Gaps 4;

QY 8 LLSADAPDAGEATAYASKDLEQLRSVSSVDEL--MTVLYPEYKMYKCOLKRGKQHNH 65
 DB 691 LEPANSEARISFPAASIASIPKTSIDAMPFTVLVPHYSEKILSLR----EILR 746
 OY 66 EOANLSRT-----EETIKFA-----AAHYTELKSIDNE 96
 DB 747 EEDQLSRVTLLEYKOLYEVEMRNPFVDTKLADENDSVIGSIDNE 792

RESULT 11
 ID 048579 PRELIMINARY; PRT; 2228 AA.

AC 048579; PRELIMINARY; PRT; 2228 AA.
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE HUMAN MI-2 AUTOANTIGEN-LIKE PROTEIN (HELICASE-LIKE PROTEIN).
 GN T19K24.8.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RC Rounsley S.D., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C.,
 RA Sykes S.M., Kaul S., Mason T.M., Kerlavage A.R., Adams M.D.,
 RA Somerville C.R., Venter J.C.;
 RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RC MEDLINE=99156233; PubMed=10048488;
 RA Asanizu E., Sato S., Kaneko T., Nakamura Y., Kotani H., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VIII.
 RT Sequence features of the regions of 1,081,958 bp covered by seventeen
 RT physically assigned P1 and TAC clones.";
 RT DNA Res. 5:379-391(1998).
 CC -1- SIMILARITY: TO HELICASE C-TERMINAL DOMAIN.
 CC EMBL: AC002342; AAC79140.1;
 DR EMBL: AB016874; BAB08833.1;
 DR InterPro: IPR000330;
 DR InterPro: IPR000953;
 DR InterPro: IPR001005;
 DR InterPro: IPR001650;
 DR InterPro: IPR001965;
 DR Pfam: PF00176; SNF2_N; 1.
 DR Pfam: PF00271; helicase_C; 1.
 DR Pfam: PF00385; chromo; 1.
 DR Pfam: PF00628; PHD; 1.
 DR PROSITE: PS50013; CHROMO_2; 2.
 DR PROSITE: PS50090; MYB_3; 1.
 KW ATP-binding; Helicase.
 SQ SEQUENCE 2228 AA; 248522 MW; 88CDBF3282C40065 CRC64;

Query Match 13.1%; Score 70.5; DB 10; Length 2228;
 Best Local Similarity 25.3%; Pred. No. 1.7e+02;
 Matches 24; Conservative 18; Mismatches 46; Indels 7; Gaps 3;
 OY 9 LLSADAPDAGEATAYASKDLEQLRSVSSVDEL--TVLYPEYKMYKCOLKRGKQHNH 66

DB 501 LEPANSEARISFPAASIASIPKTSIDAMPFTVLVPHYSEKILSLR----EILR 746
 OY 67 QANLSRTTEETIKFAAHYTELKSIDNEWRKQ 101
 DB 559 LKGLAKRKLLENYK--AKYGTAVINICEDKWKOPQ 590

RESULT 12

ID 0902H1 PRELIMINARY; PRT; 4488 AA.

AC 0902H1; PRELIMINARY; PRT; 4488 AA.
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE LEFT-RIGHT DYNEIN.
 GN DNACH11 OR LRD.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Supp D.M., Brueckner M., Kuehn M.R., Witte D.P., Lowe L.A.,
 RA McGrath J., Corrales J.M., Potter S.S.;
 RT "Targeted deletion of the ATP binding domain of left-right dynein
 RT confirms its role in specifying development of left-right
 RT asymmetries.";
 RL Development 0:0-0(1999).
 DR EMBL: AF183144; AF07922.1;
 DR MGI: MGI:1100864; Dnach11.
 DR InterPro: IPR000169;
 DR InterPro: IPR001680;
 DR InterPro: IPR001993;
 DR InterPro: IPR003593;
 DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
 DR PROSITE: PS00639; TYROL_PROTEASE_HIS; UNKNOWN_1.
 DR PROSITE: PS00678; WD_REPEATS; UNKNOWN_1.
 DR SMART: SM00382; AAA; 1.
 SQ SEQUENCE 4488 AA; 516184 MW; 1C5E0050928D949A CRC64;

Query Match 13.1%; Score 70.5; DB 11; Length 4488;
 Best Local Similarity 26.3%; Pred. No. 3.9e+02;
 Matches 26; Conservative 15; Mismatches 21; Indels 37; Gaps 5;

QY 7 LLSADAPDAGEATAYASKDLEQLRSVSSVDELMTVLYPEYKMYKCOLKRGKQ 62
 DB 1263 LKANKOLELA-----LEEMQMONSARLEFVALPEYKMKCHROEIRLLKGLW- 1311
 OY 63 HNRQANLSRTTEETIKFAAHYTELKSIDNEWRKQ 101
 DB 1312 -----DVITY-----VRSIDN-WTEQ 1328

RESULT 13

ID 056080 PRELIMINARY; PRT; 371 AA.

AC 056080; PRELIMINARY; PRT; 371 AA.
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE TLPA.
 GN TLPA.
 OS Salmomella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmomella.
 OX NCBI_Taxid=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=92291112; PubMed=1601892;
 RA Koski P., Saarialhti H., Sukupolvi S., Taira S., Rikonen P.,
 RA Osterlund K., Hurme R., Rhen M.;

RT "A new alpha-helical coiled coil protein encoded by the Salmonella
 RT typhimurium virulence plasmid."
 RL J. Biol. Chem. 267:12258-12265(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94193768; PubMed-8144657;
 RA Hume R., Namork E., Nurmiaho-Lassila E.L., Rhen M.;
 RT "Intermediate filament-like network formed in vitro by a bacterial
 RT coiled coil protein."
 RL J. Biol. Chem. 269:10675-10682(1994).
 DR EMBL; M88208; AAA4964.1; -
 SQ SEQUENCE 371 AA; 41523 MW; E0248FE76D9B00F2 CRC64;

Query Match 13.0%; Score 70; DB 2; Length 371;
 Best Local Similarity 31.2%; Pred. No. 22;
 Matches 24; Conservative 12; Mismatches 27; Indels 14; Gaps 3;

OY 16 AGEATAVASKDLEEQRSVSDVDELMTVLYPEYWKM-----YKCOLRKGWQHNRQANL 70
 DB 109 AGEQTQAARELADAAQYVDLEKDELQDRYDSLTLALSESRLRQ---QHVDYMAQL 165
 OY 71 NSRTEETIKFAAHYNT 87
 DB 166 KER-----LAAAEENT 176

RESULT 14
 ID 09E015 PRELIMINARY; PRT; 371 AA.
 AC 09E015;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
 DE TLPa.
 GN TLPa.
 OS Salmonella enterica subsp. enterica serovar Choleraesuis.
 OG Plasmid 50k virulence.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxId-119912;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RF-1;
 RA Okada N., Haneda T.;
 RT "50 kb virulence plasmid of Salmonella enterica serovar
 RT Choleraesuis."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB040415; BAB20549.1; -
 KW Plasmid.
 SQ SEQUENCE 371 AA; 41479 MW; 90548FE76D9B05A2 CRC64;

Query Match 13.0%; Score 70; DB 2; Length 371;
 Best Local Similarity 31.2%; Pred. No. 22;
 Matches 24; Conservative 12; Mismatches 27; Indels 14; Gaps 3;

OY 16 AGEATAVASKDLEEQRSVSDVDELMTVLYPEYWKM-----YKCOLRKGWQHNRQANL 70
 DB 109 AGEQTQAARELADAAQYVDLEKDELQDRYDSLTLALSESRLRQ---QHVDYMAQL 165
 OY 71 NSRTEETIKFAAHYNT 87
 DB 166 KER-----LAAAEENT 176

RESULT 15
 ID 005206 PRELIMINARY; PRT; 124 AA.
 AC 005206;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, last sequence update)
 DT 01-JAN-1998 (TREMBLrel. 05, last annotation update)

DE M PROTEIN (FRAGMENT).
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxId-1314;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ST14935;
 RA Beall B.W.;
 RT Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U92492; AAB51153.1; -
 FT NON_TER 1 124
 FT 124
 SQ SEQUENCE 124 AA; 14716 MW; 822B6A11B6F9F6ED CRC64;

Query Match 12.8%; Score 69; DB 2; Length 124;
 Best Local Similarity 25.5%; Pred. No. 7.6;
 Matches 28; Conservative 15; Mismatches 37; Indels 30; Gaps 5;

OY 1 AAFESGLDSDAEPDAGEATATASDLEEQRSVSDVDELMTVLYPEYWKM---KCOLR 57
 DB 12 AGFANQTEVKAAGPPSSPQNVSSDRDIYE-----LHEELWKEYDILKEKLD 57
 OY 58 KGWQHNRQANLIN-----SRTEETIKFAAHYNTILKSIDNEMRKTO 101
 DB 58 KD--QEEREKIELNYLKKLDKNQEREKLT-----ELDYLKIDHHEKHEQ 100

Search completed: October 17, 2001, 14:50:51
 Job time: 417 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:46:01 ; Search time 115.93 Seconds
(without alignments)
37.128 Million cell updates/sec

Title: US-09-427-657-2_COPY_32_102
Perfect score: 375
Sequence: 1 FESGDLSDAEPDAGEATAY.....CQLRKGWQHNRQANLSR 71

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues
Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.GeneSeq_0601.*
1: /SIDS8/gcgdata/geneSeq/geneSeq/AA1980.DAT.*
2: /SIDS8/gcgdata/geneSeq/geneSeq/AA1981.DAT.*
3: /SIDS8/gcgdata/geneSeq/geneSeq/AA1982.DAT.*
4: /SIDS8/gcgdata/geneSeq/geneSeq/AA1983.DAT.*
5: /SIDS8/gcgdata/geneSeq/geneSeq/AA1984.DAT.*
6: /SIDS8/gcgdata/geneSeq/geneSeq/AA1985.DAT.*
7: /SIDS8/gcgdata/geneSeq/geneSeq/AA1986.DAT.*
8: /SIDS8/gcgdata/geneSeq/geneSeq/AA1987.DAT.*
9: /SIDS8/gcgdata/geneSeq/geneSeq/AA1988.DAT.*
10: /SIDS8/gcgdata/geneSeq/geneSeq/AA1989.DAT.*
11: /SIDS8/gcgdata/geneSeq/geneSeq/AA1990.DAT.*
12: /SIDS8/gcgdata/geneSeq/geneSeq/AA1991.DAT.*
13: /SIDS8/gcgdata/geneSeq/geneSeq/AA1992.DAT.*
14: /SIDS8/gcgdata/geneSeq/geneSeq/AA1993.DAT.*
15: /SIDS8/gcgdata/geneSeq/geneSeq/AA1994.DAT.*
16: /SIDS8/gcgdata/geneSeq/geneSeq/AA1995.DAT.*
17: /SIDS8/gcgdata/geneSeq/geneSeq/AA1996.DAT.*
18: /SIDS8/gcgdata/geneSeq/geneSeq/AA1997.DAT.*
19: /SIDS8/gcgdata/geneSeq/geneSeq/AA1998.DAT.*
20: /SIDS8/gcgdata/geneSeq/geneSeq/AA1999.DAT.*
21: /SIDS8/gcgdata/geneSeq/geneSeq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneSeq/geneSeq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	375	100.0	399	20	AAW6237
2	375	100.0	419	18	AAW17837
3	375	100.0	419	18	AAW00932
4	375	100.0	419	18	AAW13833
5	375	100.0	419	18	AAW11478
6	375	100.0	419	19	AAW5751
7	375	100.0	419	19	AAW5740
8	375	100.0	419	20	AAW30518
9	375	100.0	419	20	AAW22320
10	375	100.0	419	20	AAW6203
11	375	100.0	419	21	AAW10648

12	375	100.0	419	21	AAW29048
13	375	100.0	419	21	AAW97144
14	375	100.0	419	21	AAW70749
15	375	100.0	419	21	AAW70982
16	375	100.0	419	22	AAW97570
17	375	100.0	419	22	AAW37605
18	284	75.7	418	18	AAW00934
19	284	75.7	418	19	AAW5743
20	276	73.6	415	18	AAW00933
21	276	73.6	415	19	AAW5742
22	191	50.9	350	16	AAW82686
23	191	50.9	350	20	AAW30519
24	191	50.9	350	20	AAW2321
25	191	50.9	350	21	AAW97145
26	191	50.9	350	22	AAW97577
27	88	23.5	326	19	AAW4296
28	87	23.2	325	19	AAW53240
29	87	23.2	325	22	AAW97572
30	87	23.2	354	19	AAW49036
31	87	23.2	354	19	AAW53241
32	87	23.2	354	19	AAW4293
33	87	23.2	354	21	AAW10649
34	87	23.2	354	21	AAW29049
35	87	23.2	354	21	AAW70750
36	87	23.2	354	21	AAW70983
37	87	23.2	354	22	AAW97573
38	87	23.2	354	22	AAW37606
39	86	22.9	178	20	AAW08287
40	86	22.9	321	19	AAW53243
41	86	22.9	337	20	AAW08286
42	86	22.9	358	19	AAW53242
43	86	22.9	358	19	AAW44295
44	85	22.7	358	18	AAW14992
45	84	22.4	354	22	AAW70685

ALIGNMENTS

RESULT 1
AAW6237
ID AAW6237 standard; protein: 399 AA.
XX
AC AAW6237;
DT 16-FEB-1999 (first entry)
XX
DE Human VEGF-C full length sequence.
XX
KW VEGF; VRP: vascular endothelial growth factor; VEGF-related protein;
KW recombinant; truncated; gene therapy; angiogenesis; cardiac ischemia;
KW coronary; collateral vessel development; cell growth; migration; heart;
KW lower limb ischemia; stroke; peripheral vascular disease; intestine;
KW wound healing; skin; vascular permeability.
XX
OS Homo sapiens.
XX
PN W09849300-A2.
XX
PD 05-NOV-1998.
XX
PF 20-APR-1998; 98WO-US07801.
XX
PR 25-APR-1997; 97US-0842984.
XX
PA (COLL-) COLLATERAL THERAPEUTICS.
XX
PI Bohlen P;
XX WPI: 1999-009426/01.
XX
PT New truncated vascular endothelial growth factor-related protein
PT subunits - lack part of the N-terminal sequence, used to stimulate

PT angiogenesis, e.g. for treating heart disease and ischaemia
 XX
 PS Claim 5; Fig 2D; 113pp; English.
 XX
 CC The invention relates to truncated VRP (vascular endothelial growth
 CC factor (VEGF)-related protein) subunits that have at least one amino
 CC acid N-terminal to the first Cys of the core sequence deleted. Host
 CC cells transfected or transfected with expression vectors containing
 CC nucleic acids encoding the truncated VRP subunits are used to produce
 CC the truncated proteins recombinantly. The truncated VRP subunits,
 CC optionally expressed from gene therapy vectors, have in vivo and in vitro
 CC angiogenic activity and are used to stimulate angiogenesis, particularly
 CC coronary collateral vessel development in cases of cardiac ischaemia; to
 CC stimulate endothelial cell growth and migration in vitro; to treat heart
 CC disease; to treat ischaemia (e.g. cardiac, chronic coronary or chronic
 CC lower limb ischaemia; stroke and peripheral vascular disease); to promote
 CC healing of wounds (of skin or intestines), and to increase vascular
 CC permeability. Sequences AAM86234 to AAM86239 represent full length VRP
 CC sequences from which the truncated fragments are created.
 XX
 SO Sequence 399 AA:
 Query Match 100.0%; Score 375; DB 20; Length 399;
 Best Local Similarity 100.0%; Pred. No. 7.4e-40;
 Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 FESGDLSDAEPDAGEATYASKDLEOLRSVSDVDEMTVLYPEYWKYKCOLRKGMO 60
 DB 12 fesgdlisdapepdageatayaskdleeqrlrvsvsdvdeimtvlypeywkmykcqlrkxgwg 71
 OY 61 HNRQANLNSR 71
 DB 72 hnreganlnsr 82
 RESULT 2
 AAM17837
 ID AAM17837 standard; Protein; 419 AA.
 XX
 AC AAM17837;
 XX
 DT 13-JAN-1998 (first entry)
 XX
 DE Human foetal liver kinase A binding protein flk-1bp.
 XX
 KW Foetal liver kinase 1 binding protein; human; flk-1bp;
 KW receptor tyrosine kinase; angiogenesis; antibody.
 KW wound healing; tumour; therapy; antagonist; antibody.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..20
 FT /label= Sig_peptide
 FT Protein 21..419
 FT /label= Mat_protein
 FT /note= "(Claim 10)"
 FT Peptide 21..35
 FT /label= N-terminal
 FT /note= "(Claim 9)"
 FT
 FT
 PN MO9717442-A1.
 PD 15-MAY-1997.
 XX
 PF 05-NOV-1996; 96WO-US17584.
 XX
 PR 08-NOV-1995; 95US-0554374.
 XX
 PA (IMMUNEX CORP.
 XX
 PI Lyman SD;

XX
 DR WPI: 1997-281031/25.
 DR N-PSDB: AAM68811.
 XX
 PT DNA encoding a human foetal liver kinase 1 binding protein - used
 PT to treat conditions with insufficient protein, deliver agents to
 PT cells and identify antagonists to treat protein-mediated conditions
 PS
 XX Claim 1; Page 30-32; 43pp; English.
 CC This polypeptide comprises a human foetal liver kinase 1 binding
 CC protein (flk-1bp) (see AAM17837) that binds to the receptor tyrosine
 CC kinase flk-1 expressed on vascular endothelial and other cells.
 CC The mature flk-1bp can be secreted from host cells transfected with
 CC an expression vector including an isolated flk-1bp cDNA clone (see
 CC AAM68811). Flk-1bp can be used to isolate cells to which it binds,
 CC for use in studying the roles of such cells and of flk-1 in
 CC vasculogenesis and angiogenesis. Angiogenesis inhibition or
 CC increased vascularisation may be clinically desirable (e.g. to
 CC suppress solid tumour growth or in wound healing, respectively).
 CC The flk-1bp can be administered to treat conditions with defective
 CC or insufficient flk-1. Polypeptides may also act as carriers to
 CC deliver diagnostic/therapeutic agents to cells to which flk-1bp
 CC binds, to generate antibodies, and to identify flk-1bp antagonists
 CC useful for treating flk-1bp mediated conditions.
 SO Sequence 419 AA:
 Query Match 100.0%; Score 375; DB 18; Length 419;
 Best Local Similarity 100.0%; Pred. No. 7.9e-40;
 Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 FESGDLSDAEPDAGEATYASKDLEOLRSVSDVDEMTVLYPEYWKYKCOLRKGMO 60
 DB 32 fesgdlisdapepdageatayaskdleeqrlrvsvsdvdeimtvlypeywkmykcqlrkxgwg 91
 OY 61 HNRQANLNSR 71
 DB 92 hnreganlnsr 102
 RESULT 3
 AAM00932
 ID AAM00932 standard; Protein; 419 AA.
 XX
 AC AAM00932;
 XX
 DT 10-NOV-1997 (first entry)
 XX
 DE Human Flt4 receptor tyrosine kinase ligand VEGF-C.
 XX
 KW VEGF-C; Flt4; receptor tyrosine kinase; VEGF-R-3; human;
 KW vascular endothelial growth factor receptor-3; ligand;
 KW angiogenesis; wound healing; lymph vessel; lymphangioma;
 KW cancer; metastasis; therapy; diagnosis; antibody; inhibitor.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..102
 FT /label= Prepro_peptide
 FT Peptide 32..227
 FT /note= "preferred active fragment of VEGF-C,
 FT retaining Flt4 ligand activity (Claim 15)"
 FT Peptide 103..217
 FT /note= "preferred active fragment of VEGF-C,
 FT retaining Flt4 ligand activity (Claim 12)"
 FT Peptide 103..225
 FT /note= "preferred active fragment of VEGF-C,
 FT retaining Flt4 ligand activity (Claim 13)"
 FT Peptide 103..227
 FT /note= "preferred active fragment of VEGF-C,
 FT retaining Flt4 ligand activity of VEGF-C,"
 FT Peptide 103..227
 FT /note= "preferred active fragment of VEGF-C,"

RESULT	4
AAW13833	

QY 61 HNRQANLSR 71
|||||

Db 92 hnreganlnsr 102

RESULT 5

AAW11478
ID AAW11478 standard; Protein: 419 AA.

XX AAW11478;

XX 23-APR-1997 (first entry)

XX Human vascular endothelial growth factor 2.

XX Vascular endothelial growth factor 2; VEGF2; angiogenesis;
KM endothelialisation; coronary bypass surgery; vascular graft surgery;
KW agonist; antagonist; therapy; diagnosis.

XX Homo sapiens.

XX Key Location/Qualifiers
FH Peptide 1..46
FT /label= Sig_peptide
FT 47..419FT Protein /label= Mat_protein
FT /note= "the mature protein is separately claimed
FT (Claim 5)"

XX MO9639515-A1.

XX 12-DEC-1996.

XX 06-JUN-1996; 96WO-US09001.

XX 06-JUN-1995; 95US-0465968.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Cao L, Hu J, Rosen CA;

XX WPI: 1997-043137/04.

XX N-PSDB: AAT51371.

XX DNA encoding human vascular endothelial growth factor 2 - used to
XX promote angiogenesis or endothelialisation in vascular graft surgery
XX
XX Claim 1; Fig 2; 74pp; English.XX Human vascular endothelial growth factor 2 (VEGF2) (AAW11478) is
XX structurally related to the VEGF/PDGF family and is a potent
XX mitogen for vascular endothelial cells, stimulating their growth
XX and angiogenesis. The amino acid sequence of VEGF2 was deduced
XX from a cDNA clone (AAT51371) obtd. from an early stage human (week 9)
XX embryo cDNA library. VEGF2 polypeptides can be produced in
XX transformed host cells and used to promote angiogenesis e.g. to
XX stimulate the growth of transplanted tissue following coronary
XX bypass surgery, or to promote endothelialisation in vascular graft
XX surgery. It can also be used to screen for antagonists (useful
XX e.g. for tumour therapy) and agonists of VEGF2 activity.

XX Sequence 419 AA:

Query Match 100.0%; Score 375; DB 18; Length 419;
Best Local Similarity 100.0%; Pred. No. 7.9e-40;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;OY 1 FESGDLSDAEPDAGEATYASKDLLEQLRSVSVDELMTVLYPEYWKYKCOLRRGQM 60
DB 32 fessgldlsdaepdageatayaskdleeqlrsvsvdelmtvlypeywkycqlrkxgqw 91OY 61 HNREQANLNSR 71
DB 92 hnreganlnsr 102

RESULT 6

AAW75751
ID AAW75751 standard; Protein: 419 AA.

XX AAW75751;

XX 14-DEC-1998 (first entry)

XX Vascular endothelial growth factor C protein analogue.

XX Flt4; vascular endothelial growth factor C; vascular endothelial cell;
KM lymphatic endothelial cell; myelopoiesis; angiogenesis; inflammation;
KW lymphangiogenesis; oedema; elephantiasis; Milroy's disease.

XX Synthetic.

XX Key Location/Qualifiers
FH Modified-site 156
FT /note= "Xaa can be anything other than cysteine, or
FT can be nothing"

XX WO98333917-A1.

XX 06-AUG-1998.

XX 02-FEB-1998; 98WO-US01973.

XX 05-FEB-1997; 97US-0795430.

XX (LUDM-) LUDWIG INST CANCER RES.

XX (UYHE-) UNIV HELSINKI LICENSING LTD.

XX Alitalo K, Joukov V;

XX WPI: 1998-437470/37.

XX New isolated vascular endothelial growth factor polypeptide(s) -
XX used to develop products for treating, e.g. cancers, inflammation,
XX oedema, granulocytopenia or for wound healing or tissue
XX transplantation
XX
XX Example 35; Page 143-145; 177pp; English.XX The vascular endothelial growth factor C (VEGF-C) polypeptides have
XX activities affecting growth and migration of vascular endothelial cells,
XX promoting growth of lymphatic endothelial cells and lymphatic vessels,
XX increasing vascular permeability, and affecting myelopoiesis. The
XX products can be used for stimulating angiogenesis, for inhibiting
XX angiogenesis, for stimulating lymphangiogenesis, treatment or prevention
XX of inflammation, oedema, elephantiasis, or Milroy's disease. They can
XX also be used to modulate myelopoiesis, e.g. treating granulocytopenia.
XX They can also be used for modulating the growth of endothelial cells.
XX They can also be used to stimulate lymphocyte production and maturation,
XX and to promote or inhibit trafficking of leucocytes between tissues and
XX lymphatic vessels or to affect migration in and out of the thymus.

XX Sequence 419 AA:

Query Match 100.0%; Score 375; DB 19; Length 419;
Best Local Similarity 100.0%; Pred. No. 7.9e-40;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;OY 1 FESGDLSDAEPDAGEATYASKDLLEQLRSVSVDELMTVLYPEYWKYKCOLRRGQM 60
DB 32 fessgldlsdaepdageatayaskdleeqlrsvsvdelmtvlypeywkycqlrkxgqw 91OY 61 HNREQANLNSR 71
DB 92 hnreganlnsr 102

RESULT 7
AAW5740
ID AAW5740 standard; Protein: 419 AA.
XX
AC AAW5740;
XX
DE 20-NOV-1998 (first entry)
XX
DE Human vascular endothelial growth factor C protein.
XX
KW Flt4; vascular endothelial growth factor C; vascular endothelial cell;
KW lymphatic endothelial cell; myelopoiesis; angiogenesis; inflammation;
KW lymphangiogenesis; oedema; elephantiasis; Milroy's disease.
XX
OS Homo sapiens.
XX
PN WO9833917-A1.
XX
PD 06-AUG-1998.
XX
PF 02-FEB-1998; 98WO-US01973.
XX
PR 05-FEB-1997; 97US-0795430.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
PA (UYHE-) UNIV HELSINKI LICENSING LTD.
XX
PI Alitalo K, Joukov V;
XX
DR WPI: 1998-437470/37.
DR N-PSDB; AAW52576.
XX
PT New isolated vascular endothelial growth factor polypeptide(s) -
PT used to develop products for treating, e.g. cancers, inflammation,
PT oedema, granulocytopenia or for wound healing or tissue
PT transplantation
XX
PS Claim 1; Page 112-115; 177pp; English.
XX
SQ The vascular endothelial growth factor C (VEGF-C) polypeptides have
CC activities affecting growth and migration of vascular endothelial cells,
CC promoting growth of lymphatic endothelial cells and lymphatic vessels,
CC increasing vascular permeability, and affecting myelopoiesis. The
CC products can be used for stimulating angiogenesis, for inhibiting
CC angiogenesis, for stimulating lymphangiogenesis, treatment or prevention
CC of inflammation, oedema, elephantiasis, or Milroy's disease. They can
CC also be used to modulate myelopoiesis, e.g. treating granulocytopenia.
CC They can also be used for modulating the growth of endothelial cells.
CC They can also be used to stimulate lymphocyte production and maturation,
CC and to promote or inhibit trafficking of leucocytes between tissues and
CC lymphatic vessels or to affect migration in and out of the thymus.
XX
SQ Sequence 419 AA:

Query Match 100.0%; Score 375; DB 19; Length 419;
Best Local Similarity 100.0%; Pred. No. 7.9e-40;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FESGIDLSAEPDAGEATAYASKDLSEQLRSVSDVDEMTVTVPYKWKYKCOLRKGWQ 60
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
32 fesgidlsdaepdageatayaskdlseqlrsvsvdelmtvtlypeywmkykcqikrgwq 91
OY 61 HNRQANLNSR 71
DB ||||||||||||
92 hnrqanlnsr 102

RESULT 8
AAV30518
ID AAV30518 standard; Protein: 419 AA.
XX

AC AAV30518;
XX
DE 16-NOV-1999 (first entry)
XX
DE Vascular endothelial growth factor-2 (VEGF-2).
XX
KW Human vascular endothelial growth factor-2; VEGF-2;
KW angiogenesis; blood pressure; blood flow; immune system disorder;
KW immune cell; cancer; autoimmune disorder; blood protein disorder;
KW ataxia telangiectasia; common variable immunodeficiency;
KW leukocyte adhesion deficiency syndrome; HTLV-BLV infection;
KW phagocyte bactericidal dysfunction; severe combined immunodeficiency;
KW Wiskott-Aldrich disorder; anemia; thrombocytopenia; hemoglobinuria;
KW allergy; asthma; allergic asthma.
XX
OS Homo sapiens.
XX
PN WO9946364-A1.
XX
PD 16-SEP-1999.
XX
PF 10-MAR-1999; 99WO-US05021.
XX
PR 13-MAR-1998; 98US-0042105.
PR 30-JUN-1998; 98US-0107997.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Cao L, Hu J;
XX
DR WPI: 1999-551399/46.
DR N-PSDB; AAZ10523.
XX
PT New human vascular endothelial growth factor-2, used for treating, e.g.
PT immune disorders and cancers
XX
PS Claim 12; Fig 1A-E; 222pp; English.
XX
SQ The present sequence represents vascular endothelial growth factor-2
CC (VEGF-2). The VEGF-2 polypeptides have activities similar to VEGF. The
CC VEGF-2 polypeptides stimulate the growth of vascular endothelial cells,
CC stimulate endothelial cell migration, stimulate angiogenesis, decrease
CC blood pressure, and increase blood flow. The polynucleotides and
CC polypeptides can be used for preventing, treating or ameliorating a
CC medical condition. The VEGF-2 polypeptides or polynucleotides may be
CC useful in treating deficiencies or disorders of the immune system, by
CC activating or inhibiting the proliferation, differentiation or
CC mobilization (chemotaxis) of immune cells. The etiology of these immune
CC deficiencies or disorders may be genetic, somatic, such as cancer or
CC some autoimmune disorders, acquired (e.g. by chemotherapy or toxins), or
CC infectious. Examples of immunologic deficiency syndromes include blood
CC protein disorders, ataxia telangiectasia, common variable
CC immunodeficiency, digestive syndrome, HIV infection, HTLV-BLV infection,
CC leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte
CC bactericidal dysfunction, severe combined immunodeficiency (SCIDS),
CC Wiskott-Aldrich disorder, anemia, thrombocytopenia, or hemoglobinuria.
CC They can also be used to modulate emostatic or thrombolytic activity.
CC Similarly allergic reactions and conditions such as asthma (particularly
CC allergic asthma) or other respiratory problems, may also be treated.
XX
SQ Sequence 419 AA:

Query Match 100.0%; Score 375; DB 20; Length 419;
Best Local Similarity 100.0%; Pred. No. 7.9e-40;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FESGIDLSAEPDAGEATAYASKDLSEQLRSVSDVDEMTVTVPYKWKYKCOLRKGWQ 60
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
32 fesgidlsdaepdageatayaskdlseqlrsvsvdelmtvtlypeywmkykcqikrgwq 91

KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
KM venous sore; diabetic ulcer; burns; skin graft growth; VEGC.
XX
XX Homo sapiens.
OS
PN WO200037641-A2.
PD
XX 29-JUN-2000.
XX
PF 21-DEC-1999; 99WO-US30503.
XX
XX 22-DEC-1998; 98GB-0028377.
PR 18-MAR-1999; 99US-0124967.
PR 08-NOV-1999; 99US-0164131.
XX
XX (JANC) JANSSEN PHARM NV.
PA
PI Gordon RD, Sprengel JJ, Von JR, Dijkmans JTH, Gosiowska A;
PI Dhanaraj SN, Xu J;
PI
DR WPI: 2000-442669/38.
XX
XX
PT New vascular endothelial growth factor protein, useful for treating or
PT preventing diseases associated with inappropriate angiogenesis activity
PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
XX
XX
PS Disclosure: Fig 11; 127pp; English.
XX
XX This invention describes a novel vascular endothelial growth factor-X
CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
CC vulnerrary, cytosatic, antirheumatic, antiarthritic, antipsoriatic and
CC antidiabetic activity and acts as an angiogenesis and vascularization
CC regulator. An antisense molecule of the invention is useful for treating
CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
CC retinopathy by inhibiting angiogenic activity or inappropriate
CC vascularization including formation and proliferation of new blood
CC vessels, growth and development of tissues, tissue regeneration and organ
CC and tissue repair in a subject. The products of the invention are useful
CC for preparing medicaments for treating wounds such as dermal ulcers,
CC pressure sores, venous sores, diabetic ulcers and burns and to promote
CC skin graft growth, tissue repair, proliferation of new blood vessels,
CC tissue regeneration and organ repair by promoting angiogenic activity or
CC vascularization. This sequence represents the human VEGC protein used
CC to illustrate the method of the invention.
XX
XX Sequence 419 AA:
SQ

Query Match 100.0%; Score 375; DB 21; Length 419;
Best Local Similarity 100.0%; Pred. No. 7.9e-40;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FESGLDLSAEPDAGATAYASKDEQLRSVSDLMFTVLYPEYKMYKCOLRRGQ 60
DB 32 fessgldlsdaepdagatayaskdleeqlrsvsdeltmvtlypeywmkycqlrkgywq 91
OY 61 HNRQANLNSR 71
DB 92 hnrqanlnsr 102

RESULT 12
AAB29048
ID AAB29048 standard; Protein: 419 AA.
XX
AC AAB29048;
XX
DT 31-JAN-2001 (first entry)
XX
DE Human VEGF-C protein sequence.
XX
XX Human; Flt4; fms-like tyrosine kinase 4; lymphoedema;
KW vascular endothelial growth factor receptor 3; VEGFR-3;

KW Milroy-Nonne syndrome; lymphoedema praecox; VEGF-C;
KM vascular endothelial growth factor C.
XX
XX Homo sapiens.
OS
PN WO200058511-A1.
PD
XX 05-OCT-2000.
XX
XX 26-MAR-1999; 99WO-US06133.
XX
XX 26-MAR-1999; 99WO-US06133.
PR 26-MAR-1999; 99WO-US06133.
XX
XX (LUDMIG INST CANCER RES.
PA (UYHE-) UNIV HELSINKI LICENSING LTD OY.
PA (UYPI-) UNIV PITTSBURGH.
XX
XX Ferrell RE, Allitalo K, Finegold DN, Karkkainen M;
PI
PI
XX
XX WPI: 2000-679298/66.
DR N-PSDB; AAC62406.
XX
XX
PT Screening a human subject for increased risk of developing a lymphatic
PT disorder, comprises assaying a nucleic acid to determine a mutation
PT altering the sequence of a vascular endothelial growth factor
PT receptor-3 -
XX
XX
PS Disclosure: Page 60-61; 76pp; English.
XX
XX The present sequence is the protein sequence for the human vascular
CC endothelial growth factor C (VEGF-C). It was used to demonstrate the
CC methods of the invention, which involve the screening of individuals to
CC determine which vascular endothelial growth factor receptor 3 (VEGFR-3,
CC also known as Flt4 or fms-like tyrosine kinase 4) alleles they possess
CC and thus their likelihood of developing hereditary lymphoedema.
CC Conditions associated with lymphoedema include Milroy-Nonne syndrome,
CC which is early onset lymphoedema and lymphoedema praecox, which is late
CC onset.
XX
XX Sequence 419 AA:
SQ

Query Match 100.0%; Score 375; DB 21; Length 419;
Best Local Similarity 100.0%; Pred. No. 7.9e-40;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FESGLDLSAEPDAGATAYASKDEQLRSVSDLMFTVLYPEYKMYKCOLRRGQ 60
DB 32 fessgldlsdaepdagatayaskdleeqlrsvsdeltmvtlypeywmkycqlrkgywq 91
OY 61 HNRQANLNSR 71
DB 92 hnrqanlnsr 102

RESULT 13
AA97144
ID AA97144 standard; Protein: 419 AA.
XX
AC AA97144;
XX
DT 22-DEC-2000 (first entry)
XX
DE Vascular endothelial growth factor-2 (VEGF-2).
XX
XX Vascular endothelial growth factor 2; VEGF-2; retina; angiogenesis;
KW treatment; injury; degeneration; photoreceptors; eye;
KW angiod streaks; retinitis; pigmentosa; human;
KW age-related macular degeneration; diabetic retinopathy.
XX
XX Homo sapiens.
OS
XX W0200045835-A1.

```

XX 10-AUG-2000.
PD 07-FEB-2000; 2000MO-US03047.
XX
XX 08-FEB-1999; 99US-0119179.
XX 12-FEB-1999; 99US-0119926.
XX 03-JUN-1999; 99US-0137796.
XX 22-DEC-1999; 99US-0171505.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Alderson R, Weider R, Roschke V, Ruben SM;
XX WPI; 2000-532862/48.
XX N-PSDB; AAA52080.
XX
XX Treating injury or degeneration of photoreceptors comprises
XX administering to a subject vascular endothelial growth factor 2
XX (VEGF-2)
XX
XX Claim 31; Fig 1a-e; 252pp; English.
XX
XX Administration of vascular endothelial growth factor 2 (VEGF-2)
XX to a patient can be used for treating injury or degeneration of
XX photoreceptors associated with e.g. angiod streaks, retinitis
XX pigmentosa, age-related macular degeneration, diabetic retinopathy,
XX etc. VEGF-2 promotes angiogenesis, the formation of new blood
XX vessels in the retina.
XX
XX Sequence 419 AA:
XX
XX Query Match
XX Best Local Similarity 100.0%; Score 375; DB 21; Length 419;
XX Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 FESGIDLSDAEPDAGEATAYASKDLBEOILRSVSDVDEMTVLYPEYWKYKCOLRKGQW 60
XX 32 fsgldlsdaepdageatayaskdlbEOILRSVSDVDEMTVLYPEYWKYKCOLRKGQW 91
XX
XX 61 HNRQANINSR 71
XX 92 hnreqanlnsr 102
XX
XX Db
XX
XX RESULT 14
XX AAY70749
XX ID AAY70749 standard; Protein: 419 AA.
XX
XX AAY70749:
XX
XX 17-AUG-2000 (first entry)
XX
XX Human prepro-vascular endothelial growth factor C.
XX
XX Human; receptor tyrosine kinase; RTK; Flt4; fms-like tyrosine kinase 4;
XX VEGFR-3; vascular endothelial growth factor receptor-3; chromosome 5q35;
XX cytostatic; tumour imaging; anti-tumour therapy; treatment; diagnosis;
XX neoplastic disease; lymphoma; carcinoma; breast; squamous cell; melanoma;
XX sarcoma; malignancy; VEGF-C; vascular endothelial growth factor C.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..31
XX Peptide /label- Signal_peptide
XX Peptide 32..103
XX /label- N-terminal_peptide
XX /note- "cleavage of this peptide from partially processed
XX VEGF-C produces a fully processed mature form of VEGF-C
XX of 21-23 kD which has high affinity to VEGFR-2"
XX 104..227
XX Protein

```

```

XX Peptide
XX 228..419
XX /label- Mature_VEGF-C
XX /label- C-terminal_peptide
XX /note- "Has a pattern of spaced cysteine residues
XX reminiscent of a Baldian ring 3 protein (BR3P) sequence;
XX cleavage of signal peptide and the C-terminal
XX peptide produces a partially processed form of VEGF-C of
XX about 29 kD which has high affinity to Flt4 (VEGFR-3)"
XX 113..213
XX /note- "binds and stimulates VEGF-C receptors; Cys
XX at position 156 is essential for VEGFR-2 binding and at
XX 165 is essential for VEGFR-2 and VEGFR-3 binding"
XX 131..211
XX /note- "important for VEGF-C activity"
XX
XX WO200021560-A1.
XX
XX 20-APR-2000.
XX
XX 08-OCT-1999; 99WO-US23525.
XX
XX 09-OCT-1998; 98US-0169079.
XX
XX (LUDWIG INST CANCER RES.
XX (UDW-) UNIV HELSINKI LICENSING LTD OY.
XX
XX Altalo K, Kaipinen A, Valltoia R, Jusella L;
XX WPI; 2000-317850/27.
XX
XX Treating neoplastic diseases such as lymphoma, carcinomas, melanomas
XX and sarcomas, involves administering a compound capable of inhibiting
XX binding of ligand proteins to fms-like tyrosine kinase-4 receptor
XX
XX Example 15-17; Page 140-142; 148pp; English.
XX
XX The patent discloses a method to treat neoplastic disease characterised
XX by expression of fms-like tyrosine kinase 4 (Flt4) receptor (also
XX referred as vascular endothelial growth factor receptor-3, VEGFR-3) in
XX endothelial cells of blood vessels adjacent to malignant neoplasm. The
XX method involves administering a compound that inhibits binding of a
XX ligand to Flt4 thereby inhibiting Flt4 mediated proliferation of vascular
XX endothelial cells. The compound is useful for treating neoplastic disease
XX such as breast carcinomas, squamous cell carcinomas, lymphomas, melanomas
XX and sarcomas. Flt4 receptor tyrosine kinase binding compounds can be used
XX for manufacturing medicament useful for diagnostic screening, imaging and
XX treatment of malignancies characterised by Flt4-expressing blood cells.
XX The Flt4 gene maps to chromosomal region 5q35 and is expressed as 5.8 kb
XX and 4.5 kb mRNAs which differ in their 3' sequences and are
XX differentially expressed in HEL and DAMI cell lines. Flt4
XX belongs to a subfamily of class III receptor tyrosine kinases (RTKs).
XX It is used as a target for tumour imaging and anti-tumour therapy.
XX The present sequence is a human prepro-vascular endothelial growth
XX factor C (VEGF-C), a specific example of Flt4 binding compound.
XX
XX Sequence 419 AA:
XX
XX Query Match
XX Best Local Similarity 100.0%; Score 375; DB 21; Length 419;
XX Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 FESGIDLSDAEPDAGEATAYASKDLBEOILRSVSDVDEMTVLYPEYWKYKCOLRKGQW 60
XX 32 fsgldlsdaepdageatayaskdlbEOILRSVSDVDEMTVLYPEYWKYKCOLRKGQW 91
XX
XX 61 HNRQANINSR 71
XX 92 hnreqanlnsr 102
XX
XX Db
XX
XX RESULT 15
XX AAY70982

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:47:10 ; Search time 62.93 Seconds
(without alignments)
23.231 Million cell updates/sec

Title: US-09-427-657-2_COPY_32_102
Perfect score: 375
Sequence: 1 FESGLDSDAEPAGEATAY.....COLRKGGMQHNREGANLNSR 71

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCIOS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	375	100.0	419	2	US-08-999-811-2
2	375	100.0	419	3	US-09-042-105-2
3	375	100.0	419	3	US-09-042-105-18
4	375	100.0	419	4	US-08-795-430-8
5	375	100.0	419	4	US-08-510-133A-35
6	369	98.4	419	5	PCR-US96-09001-2
7	284	75.7	418	4	US-08-795-430-13
8	276	73.6	415	4	US-08-795-430-11
9	191	50.9	350	2	US-08-999-811-4
10	191	50.9	350	2	US-08-824-996-2
11	191	50.9	350	3	US-09-042-105-4
12	191	50.9	350	4	US-08-510-133A-33
13	191	50.9	350	4	US-08-585-895-33
14	87	23.2	325	4	US-08-915-795-3
15	87	23.2	354	4	US-08-915-795-5
16	86	22.9	321	4	US-08-915-795-9
17	86	22.9	358	4	US-08-915-795-8
18	64	17.1	109	2	US-08-070-116A-4
19	64	17.1	110	3	US-08-444-644-44
20	64	17.1	326	2	US-08-808-720-3
21	64	17.1	327	2	US-08-761-277A-47
22	64	17.1	328	3	US-08-808-720-1
23	64	17.1	331	3	US-08-808-720-5
24	64	17.1	331	3	US-08-808-720-7
25	64	17.1	382	1	US-08-470-299-7
26	64	17.1	382	1	US-08-470-299-10
27	64	17.1	443	5	PCM-US96-13152-4

28	64	17.1	467	1	US-08-704-744-81	Sequence 81, Appl
29	64	17.1	467	2	US-07-916-098A-45	Sequence 45, Appl
30	64	17.1	467	4	US-08-523-894-8	Sequence 8, Appl
31	64	17.1	467	4	US-08-523-894-10	Sequence 10, Appl
32	64	17.1	467	4	US-08-523-894-12	Sequence 12, Appl
33	63	16.8	147	4	US-08-905-223-358	Sequence 358, App
34	61	16.3	911	2	US-08-484-438-10	Sequence 10, Appl
35	61	16.3	1220	3	US-08-930-998A-2	Sequence 2, Appl
36	60	16.0	109	3	US-08-444-644-30	Sequence 30, Appl
37	60	16.0	326	2	US-08-656-586-9	Sequence 2, Appl
38	60	16.0	432	3	US-08-477-460B-2	Sequence 2, Appl
39	60	16.0	432	3	US-08-379-516-2	Sequence 2, Appl
40	60	16.0	432	4	US-09-329-916-2	Sequence 2, Appl
41	60	16.0	432	4	US-08-485-372A-2	Sequence 2, Appl
42	60	16.0	432	5	PCR-US93-07422-2	Sequence 2, Appl
43	60	16.0	442	1	US-08-461-968A-5	Sequence 5, Appl
44	60	16.0	442	2	US-08-462-571-5	Sequence 5, Appl
45	60	16.0	450	2	US-08-788-800-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-999-811-2
; Sequence 2, Application US/08999811
; Patent No. 5932540
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: CAO, LIANG
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESS: STERNE, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/999, 811
; FILING DATE: HEREMITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/207, 550
; FILING DATE: 8-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465, 968
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: MARKOWITZ, KAREN R.
; REGISTRATION NUMBER: 36,351
; REFERENCE/DOCKET NUMBER: 1488.1000004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-999-811-2
Query Match 100.0%; Score 375; DB 2; Length 419;
Best Local Similarity 100.0%; Pred. No. 2.2e-42;

Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FESGDLSDAEPDAGEATAYASKDLRVSVDLMTVLYPEYWKMKCOLRKGQ 60
|||||
DB 32 FESGDLSDAEPDAGEATAYASKDLRVSVDLMTVLYPEYWKMKCOLRKGQ 91

QY 61 HNRQANLNSR 71
|||||
DB 92 HNRQANLNSR 102

RESULT 2
US-09-042-105-2
; Sequence 2, Application US/09042105
; Patent No. 6040157
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: CAO, LIANG
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,105
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/207,550
; FILING DATE: 8-MAR-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,968
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TO BE ASSIGNED
; FILING DATE: 24-DEC-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: ERIC K. STEFFE
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-042-105-2

Query Match 100.0%; Score 375; DB 3; Length 419;
Best Local Similarity 100.0%; Pred. No. 2.2e-42;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FESGDLSDAEPDAGEATAYASKDLRVSVDLMTVLYPEYWKMKCOLRKGQ 60
|||||
DB 32 FESGDLSDAEPDAGEATAYASKDLRVSVDLMTVLYPEYWKMKCOLRKGQ 91

QY 61 HNRQANLNSR 71
|||||
DB 92 HNRQANLNSR 102

RESULT 3
US-09-042-105-18
; Sequence 18, Application US/09042105
; Patent No. 6040157
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: CAO, LIANG
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,105
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/207,550
; FILING DATE: 8-MAR-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,968
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TO BE ASSIGNED
; FILING DATE: 24-DEC-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: ERIC K. STEFFE
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-042-105-18

Query Match 100.0%; Score 375; DB 3; Length 419;
Best Local Similarity 100.0%; Pred. No. 2.2e-42;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FESGDLSDAEPDAGEATAYASKDLRVSVDLMTVLYPEYWKMKCOLRKGQ 60
|||||
DB 32 FESGDLSDAEPDAGEATAYASKDLRVSVDLMTVLYPEYWKMKCOLRKGQ 91

QY 61 HNRQANLNSR 71
|||||
DB 92 HNRQANLNSR 102

RESULT 4
US-08-795-430-8
; Sequence 8, Application US/08795430
; Patent No. 6130071
; GENERAL INFORMATION:
; APPLICANT: Aitalo, Kari
; APPLICANT: Joukov, Vladimir
; TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
; TITLE OF INVENTION: Protein and Gene, Mutants Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,430
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/00427
; FILING DATE: 01-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/671,573
; FILING DATE: 28-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/601,132
; FILING DATE: 14-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/585,895
; FILING DATE: 12-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/510,133
; FILING DATE: 01-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,011
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 28967/33691
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-795-430-8

Query Match 100.0%; Score 375; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 2.2e-42;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FESGDLSDAEPDAGATAYASKDLEQLRSVSSVDLMTLVLPYWKMYKCOLRKGGMQ 60
|||||
DB 32 FESGDLSDAEPDAGATAYASKDLEQLRSVSSVDLMTLVLPYWKMYKCOLRKGGMQ 91
|||||
QY 61 HNRQANLNSR 71
|||||
DB 92 HNRQANLNSR 102

RESULT 5
US-08-510-133A-35
; Sequence 35, Application US/08510133A
; Patent No. 6221839
; GENERAL INFORMATION:
; APPLICANT: Aitalo, Kari
; APPLICANT: Joukov, Vladimir
; TITLE OF INVENTION: Receptor Ligand
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/510,133A
; FILING DATE: 01-AUG-1995
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 28113/32863
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-08-510-133A-35

Query Match 100.0%; Score 375; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 2.2e-42;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FESGDLSDAEPDAGATAYASKDLEQLRSVSSVDLMTLVLPYWKMYKCOLRKGGMQ 60
|||||
DB 32 FESGDLSDAEPDAGATAYASKDLEQLRSVSSVDLMTLVLPYWKMYKCOLRKGGMQ 91
|||||
QY 61 HNRQANLNSR 71
|||||
DB 92 HNRQANLNSR 102

RESULT 6
PCT-US96-09001-2
; Sequence 2, Application PC/TUS9609001
; GENERAL INFORMATION:
; APPLICANT: HU, ET AL.
; TITLE OF INVENTION: Human Vascular Endothelial Growth Factor 2
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILLILAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/09001
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/465,968
FILING DATE: 6 JUN 95
APPLICATION NUMBER: 08/207,550
FILING DATE: 8 MAR 1994
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-288
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
PCT-US96-09001-2

Query Match 98.4%; Score 369; DB 5; Length 419;
Best Local Similarity 98.6%; Pred. No. 1.4e-41;
Matches 70; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FESGGLSDAEPAGATAYASKDLEQLRSVSSVDELMVTLYPEYWKMKCOLRGGWQ 60
DB 32 FESGGLSDAEPAGATAYASKDLEQLRSVSSVDELMVTLYPEYWKMKCOLRGGWQ 91
QY 61 HNREQANLNSR 71
DB 92 HNREQANLNSR 102

RESULT 7

US-08-795-430-13
Sequence 13, Application US/08795430
Patent No. 6130071
GENERAL INFORMATION:
APPLICANT: Aitalo, Karl
APPLICANT: Joukov, Vladimir
TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
TITLE OF INVENTION: Protein and Gene, Mutants Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,430
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/ET96/00427
FILING DATE: 01-AUG-1996
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/671,573
FILING DATE: 28-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/601,132
FILING DATE: 14-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/585,895
FILING DATE: 12-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/510,133
FILING DATE: 01-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,011
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: GASS, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28967/33691
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 418 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-795-430-13

Query Match 75.7%; Score 284; DB 4; Length 418;
Best Local Similarity 74.6%; Pred. No. 4e-30;
Matches 53; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

QY 1 FESGGLSDAEPAGATAYASKDLEQLRSVSSVDELMVTLYPEYWKMKCOLRGGWQ 60
DB 31 YESGHGYEEEPGAGPKAHASKDLEQLRSVSSVDELMVTLYPEYWKMKCOLRGGWQ 90
QY 61 HNREQANLNSR 71
DB 91 HNREHSSDTR 101

RESULT 8

US-08-795-430-11
Sequence 11, Application US/08795430
Patent No. 6130071
GENERAL INFORMATION:
APPLICANT: Aitalo, Karl
APPLICANT: Joukov, Vladimir
TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
TITLE OF INVENTION: Protein and Gene, Mutants Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,430
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/ET96/00427
FILING DATE: 01-AUG-1996
PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/671,573
FILING DATE: 28-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/601,132
FILING DATE: 14-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/585,895
FILING DATE: 12-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/510,133
FILING DATE: 01-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,011
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: GASS, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28967/33691
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 415 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-795-430-11

Query Match 73.6%; Score 276; DB 4; Length 415;
Best Local Similarity 74.6%; Pred. No. 4,8e-29;
Matches 53; Conservative 5; Mismatches 9; Indels 4; Gaps 1;

QY 1 FESGDLSDAEPDAGEATVASKDLDEQLRSVSDVLTMTVLPYWKMKQQLRKGGWQ 60
||||| 1:||||| 1: |||||||||||||||:|||||:|||||:|||||
DB 32 FESGDLSDAEPDAGEATVASKDLDEQLRSVSDVLTMTVLPYWKMKQQLRKGGWQ 91

QY 61 HNRQANLSR 71
| 11:|
DB 92 ---OPTLNR 98

RESULT 9
US-08-999-811-4
Sequence 4, Application US/08999811
Patent No. 5932540
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/999,811
FILING DATE: HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550

FILING DATE: 8-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: MARKOWICZ, KAREN R.
REGISTRATION NUMBER: 36,351
REFERENCE/DOCKET NUMBER: 1488.1000004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-999-811-4

Query Match 50.9%; Score 191; DB 2; Length 350;
Best Local Similarity 100.0%; Pred. No. 1.1e-17;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 MTVLPEYWKMKQQLRKGGWQHNRQANLSR 71
||||| 1:||||| 1: |||||||||||||||
DB 1 MTVLPEYWKMKQQLRKGGWQHNRQANLSR 33

RESULT 10
US-08-824-996-2
Sequence 2, Application US/08824996B
Patent No. 5935820
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: Polynucleotides Encoding Vascular Endothelial Growth
FILE REFERENCE: P112D1
CURRENT APPLICATION NUMBER: US/08/824,996B
CURRENT FILING DATE: 1997-03-27
EARLIER APPLICATION NUMBER: 08/207,550
EARLIER FILING DATE: 1994-03-08
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 2
LENGTH: 350
TYPE: PRT
ORGANISM: Homo sapiens
US-08-824-996-2

Query Match 50.9%; Score 191; DB 2; Length 350;
Best Local Similarity 100.0%; Pred. No. 1.1e-17;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 MTVLPEYWKMKQQLRKGGWQHNRQANLSR 71
||||| 1:||||| 1: |||||||||||||||
DB 1 MTVLPEYWKMKQQLRKGGWQHNRQANLSR 33

RESULT 11
US-09-042-105-4
Sequence 4, Application US/09042105
Patent No. 6040157
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:

ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,105
FILING DATE: HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550
FILING DATE: 8-MAR-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TO BE ASSIGNED
FILING DATE: 24-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: ERIC K. STEFFE
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-042-105-4

Query Match 50.9%; Score 191; DB 3; Length 350;
Best Local Similarity 100.0%; Pred. No. 1.1e-17;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 MTVLPYWKMKCOLRKGWQHNRQANLSNR 71
DB 1 MTVLPYWKMKCOLRKGWQHNRQANLSNR 33

RESULT 12
US-08-510-133A-33
Sequence 33, Application US/08510133A
Patent No. 6221839
GENERAL INFORMATION:
APPLICANT: Allitalo, Kari
Joukov, Vladimir
TITLE OF INVENTION: Receptor Ligand
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/510,133A
FILING DATE: 01-Aug-1995
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28113/32863
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-510-133A-33

Query Match 50.9%; Score 191; DB 4; Length 350;
Best Local Similarity 100.0%; Pred. No. 1.1e-17;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 MTVLPYWKMKCOLRKGWQHNRQANLSNR 71
DB 1 MTVLPYWKMKCOLRKGWQHNRQANLSNR 33

RESULT 13
US-08-585-895-33
Sequence 33, Application US/08585895
Patent No. 6245530
GENERAL INFORMATION:
APPLICANT: Allitalo, Kari
Joukov, Vladimir
TITLE OF INVENTION: Receptor Ligand
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/585,895
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28113/33072
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-585-895-33

APPLICANT: Andrew F. WILKS

Search completed: October 17, 2001, 14:47:11
Job time: 197 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:48:40 ; Search time 78.16 Seconds
(without alignments)
69.196 Million cell updates/sec

Title: US-09-427-657-2_COPY_32_102

Perfect score: 375

Sequence: 1 FESGLDLSDAEPDAGEATAY.....COLRKGGMHNRQANLSR 71

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	375	100.0	419	2	S69207
2	69	18.4	786	2	T16509
3	67	17.9	453	2	H81151
4	67	17.9	453	2	D81870
5	66	17.6	703	2	B82148
6	65	17.3	2408	2	T24483
7	64	17.1	284	2	A75152
8	64	17.1	327	1	G4HU
9	64	17.1	1403	2	T49093
10	63.5	16.9	120	2	E71980
11	63.5	16.9	427	2	G82827
12	63.5	16.9	493	2	T29030
13	63	16.8	246	2	T46249
14	63	16.8	366	2	T46249
15	62.5	16.7	614	2	T42649
16	62.5	16.7	1040	2	G81849
17	62	16.5	210	2	G81849
18	62	16.5	210	2	A81092
19	62	16.5	371	2	A44122
20	62	16.5	437	2	I51238
21	62	16.5	1729	2	T43403
22	61.5	16.4	441	2	A48455
23	61	16.3	1220	2	T06403
24	60.5	16.1	474	2	A83347
25	60	16.0	304	2	S41582
26	60	16.0	326	1	G2HU
27	60	16.0	436	2	S20060
28	60	16.0	436	2	I51237
29	59	15.7	234	2	PT0207

30	59	15.7	234	2	E72001	hypothetical prote
31	59	15.7	234	2	E86622	hypothetical prote
32	59	15.7	243	2	F81536	hypothetical prote
33	59	15.7	246	2	T35934	probable NAD(P)H o
34	59	15.7	255	4	S31866	Ig gamma-1 chain C
35	59	15.7	330	1	G4HU	Ig gamma-1 chain C
36	59	15.7	374	2	S69339	Ig heavy chain V r
37	59	15.7	377	2	A60764	Ig gamma-3 chain C
38	59	15.7	377	2	A23511	Ig gamma-3 chain C
39	59	15.7	804	2	T37821	probable dna repai
40	59	15.7	843	2	D66495	unknown protein (I
41	59	15.7	1005	2	T31333	beta-galactosidase
42	59	15.7	1289	2	B72354	conserved hypotet
43	58.5	15.6	491	1	S24354	p53-binding protei
44	58.5	15.6	2488	2	T42739	guanine nucleotide
45	58	15.5	312	2	S35295	rtfb protein - Yer

ALIGNMENTS

RESULT 1
S69207
Vascular endothelial growth factor C precursor - human
N/Alternate names: FLT4 ligand DHM
C/Species: Homo sapiens (man)
C/Date: 27-Apr-1996 #sequence revision 01-Nov-1996 #text_change 08-Oct-1999
C/Accession: S69207; S61795; S71443; S69208; G02659
R/Joukov, V.; Pajusola, K.; Kaipainen, A.; Chillov, D.; Lahtinen, I.; Kukk, E.; Saksel
EMBO J. 15, 1751, 1996
A/Title: Corrigendum: A novel vascular endothelial growth factor, VEGF-C, is a ligand
A/Reference number: S69207; MUID:96203094
A/Accession: S69207
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-419 <JOU>
A/Cross-references: EMBL:X94216; NID:g1177488; PIDD:CAA63907.1; PID:e221096; PID:g118
A/Note: The nucleotide sequence was submitted to the EMBL Data Library, December 1995
A/Note: only a part of the translation is shown
A/Note: this is a revision to the sequence from reference S61795
R/Joukov, V.; Pajusola, K.; Kaipainen, A.; Chillov, D.; Lahtinen, I.; Kukk, E.; Saksel
EMBO J. 15, 290-298, 1996
A/Title: A novel vascular endothelial growth factor, VEGF-C, is a ligand for the FLT4
A/Reference number: S61795; MUID:96178224
A/Accession: S61795
A/Status: nucleic acid sequence not shown; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 70-419 <JOU1>
A/Note: this sequence has been revised in reference S69207
A/Accession: S71443
A/Molecule type: protein
A/Residues: 1-419 <JOU2>
R/Lee, J.; Gray, A.; Yuan, J.; Luoh, S.M.; Avraham, H.; Wood, W.I.
submitted to the EMBL Data Library, December 1995
A/Description: Vascular endothelial growth factor related protein (VRP): A ligand and
A/Reference number: S69208
A/Accession: S69208
A/Molecule type: mRNA
A/Residues: 1-419 <LEE>
A/Cross-references: EMBL:U43142; NID:g1150988; PIDD:AAA85214.1; PID:g1150989
R/Morris, J.C.
submitted to the EMBL Data Library, May 1996
A/Reference number: H01557
A/Accession: G02659
A/Status: preliminary; translated from GR/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-419 <MOR>
A/Cross-references: EMBL:U58111; NID:g1373426; PIDD:AA802909.1; PID:g1373427
C/Genetics:
A/Gene: GDB:VEGFC; VRP
A/Cross-references: GDB:3890883; OMIM:601528
F/1-12/Domain: signal sequence #status predicted <SIG>
F/13-102/Domain: propeptide #status predicted <PRO>

F:103-419/Product: vascular endothelial growth factor C #status experimental <MAT>

Query Match 100.0% Score 375; DB 2; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.9e-35;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FESGDLSDAEPDAGATAYASKDLEQLRSVSVDELMTVLYPEYKMKYKCOLRKGGMQ 60
DB 32 FESGDLSDAEPDAGATAYASKDLEQLRSVSVDELMTVLYPEYKMKYKCOLRKGGMQ 91

OY 61 HNRQANLNSR 71
DB 92 HNRQANLNSR 102

RESULT 2

hypothetical protein F59A6.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T16509

R:Nhan, M.

Submitted to the EMBL Data Library, December 1995

A:Description: The sequence of C. elegans cosmid F59A6.

A:Reference number: Z18526

A:Accession: T16509

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-786 <NHA>

A:Cross-references: EMBL:U41994; NID:g1123047; PID:g1123052; PIDN:AAA83456.1; CESP:F59A6

A:Gene: CESP:F59A6.3

A:Introns: 106/1; 620/1; 634/3; 675/3; 775/1

Query Match 18.4% Score 69; DB 2; Length 786;
Best Local Similarity 33.3%; Pred. No. 7.1;
Matches 18; Conservative 9; Mismatches 17; Indels 10; Gaps 2;

OY 20 YASKDLEQLRSVSVDELMTVLYPEYKMKYKCOLRKGGMQ 64
DB 719 YTSPTSSQIKSSVSGELITQVCPQSYVFETALQPKIKYK-LKTKRWAGSPE 771

RESULT 3

polyA polymerase NMB0843 [imported] - Neisseria meningitidis (strain MC58 serogroup B)

C:Species: Neisseria meningitidis

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C:Accession: H81151

R:Retelien, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizzi, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V.

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755

A:Accession: H81151

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-433 <TET>

A:Cross-references: GB:AE002437; GB:AE002098; NID:g7226072; PIDN:AAF41254.1; PID:g722607

C:Genetics:

A:Gene: NMB0843

Query Match 17.9% Score 67; DB 2; Length 453;
Best Local Similarity 27.5%; Pred. No. 6.5;
Matches 19; Conservative 16; Mismatches 20; Indels 14; Gaps 2;

OY 3 SGLDLS-----AEPDAGATAYASKDLEQLRSVSVDELMTVLYPEYK 48

DB 252 NGEFDPDDIHLNLNRVSDGIAGKMTVLTAKNTDERLRADRSVSGFVLALMPPELR 311

OY 49 MYKCOLRKG 57
DB 312 HWKSNLQOG 320

RESULT 4

DB1870

Probable polynucleotide adenylyltransferase (EC 2.7.7.19) NMA1053 [imported] - Neisse

C:Species: Neisseria meningitidis

C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C:Accession: D81870

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

; Holtroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis z2491.

A:Reference number: A81775; MUID:20222556

A:Accession: D81870

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-453 <PAR>

A:Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84318.1; PID:g737

A:Experimental source: serogroup A, strain z2491

C:Genetics:

A:Gene: pcmb, NMA1053

C:Keywords: nucleotidyltransferase

Query Match 17.9% Score 67; DB 2; Length 453;
Best Local Similarity 27.5%; Pred. No. 6.5;
Matches 19; Conservative 16; Mismatches 20; Indels 14; Gaps 2;

OY 3 SGLDLS-----AEPDAGATAYASKDLEQLRSVSVDELMTVLYPEYK 48
DB 252 NGEFDPDDIHLNLNRVSDGIAGKMTVLTAKNTDERLRADRSVSGFVLALMPPELR 311

OY 49 MYKCOLRKG 57
DB 312 HWKSNLQOG 320

RESULT 5

DB2148

ATP-dependent helicase, Ding family VC1855 [imported] - Vibrio cholerae (strain N1696

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: B82148

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.

Charidson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers

Li, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: B82148

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-703 <HET>

A:Cross-references: GB:AE004261; GB:AE003852; NID:g9656382; PIDN:AAF95003.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC1855

A:Map position: 1

Query Match 17.6% Score 66; DB 2; Length 703;
Best Local Similarity 31.6%; Pred. No. 14;
Matches 18; Conservative 9; Mismatches 20; Indels 10; Gaps 2;

OY 9 DAEPDAGATAYASKDLEQLRSVSVDELMTVLYPEYKMKYKCOLRKGGMQ 59
DB 514 EMEPDAGATAYASKDLEQLRSVSVDELMTVLYPEYKMKYKCOLRKGGMQ 566

[illegible]

```

C>Date: 02-Apr-1982 #sequence revision 02-Apr-1982 #text_change 16-Jul-1999
C:Accession: A90933; A90249; A02150
R:Ellison, J.; Buxbaum, J.; Hood, L.
DNA 1, 11-18, 1981
A>Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.
A:Reference number: A90933; MUID:83157104
A:Accession: A90933
A:Molecule type: DNA
A:Residues: 1-327 <ELL>
A>Note: The sequence was determined from the germ-line gene
R:Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.
Biochem. J. 117, 33-47, 1970
A>Title: Human immunoglobulin subclasses. Partial amino acid sequence of the constant
A:Reference number: A90249; MUID:70207560
A:Accession: A90249
A:Molecule type: Protein
A:Residues: 1-30; 81-326 <PIN>
C:Genetics:
A:Gene: GDB:IGHG4
A:Cross-references: GDB:119340; OMIM:147130
A:Map position: 14q32.33-14q32.33
A:Introns: 99/1; 111/1; 221/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (L) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Keywords: duplication; immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-85/Domain: Immunoglobulin homology <IMI>
F:99-110/Region: hinge
F:134-203/Domain: immunoglobulin homology <IM3>
F:240-307/Domain: immunoglobulin homology <IM3>
F:14/Disulfide bonds: Interchain (to light chain) #status experimental
F:27-83,141-201,247-305/Disulfide bonds: #status predicted
F:106,109/Disulfide bonds: Interchain (to heavy chain) #status experimental
F:177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match      17.1%; Score 64; DB 1; Length 327;
Best Local Similarity 29.7%; Pred. No. 9.9;
Matches 19; Conservative 11; Mismatches 24; Indels 10; Gaps 2;

OY      5 LDLSAEP-----DAGEATAAVASKDEQLSVSSVDLEMTLVLPDY--WKMYKQOL 54
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      144 VDVSQEDPEVGPNKYVDGEVHNNAKTFRPQDFNSTRYWVSVLIVLDMDWLNGKEIKCKV 203

OY      55 RKKG 58
          |
Db      204 SNKG 207

RESULT      9
T49093
hypothetical protein F4F15.250 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T49093
R:Alcazar, J.P.; Clabault, G.; Cortet, A.; Maché, R.; Mewes, H.W.; Rudd, S.; Lemcke,
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z25015
A:Accession: T49093
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1403 <ALIC>
A:Cross-references: EMBL:AL049711; GSPDB:GN00061; ATSP:F4F15.250
C:Experimental source: cultivar Columbia; BAC clone F4F15
C:Genetics:
A:Gene: ATSP:F4F15.250
A:Map position: 3
A:Introns: 103/1; 123/3; 187/1; 234/1; 286/2; 351/1; 385/1; 430/2; 485/2; 606/3; 641/

Query Match      17.1%; Score 64; DB 2; Length 1403;
Best Local Similarity 32.1%; Pred. No. 51;
Matches 23; Conservative 11; Mismatches 30; Indels 12; Gaps 4;

```

```

Oy      2  ESKDLSDAEPDAGEATAY-----ASKDLEPOLRSVSVDELMTVLV--PE--YMKWY 50
      ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      90  ESOVENNDAPKQGEIRLTPVSVKTSQSGCKMELQINPGDSVMDIROFLDAPETCTFTCY 149
      ||::||::||::||::||::||::||::||::||::||::||::||::||::||
Oy      51  KCQLR-KGGWQHNRQAN 67
      :||::||::||::||::||::||::||::||::||::||::||::||::||
Db      150  ELLRNKQDGTTHLEDDYN 167
      ||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 10
E71980
hypothetical protein jhp0051 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C:Accession: E71980
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557
A:Accession: E71980
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-120 <ARN>
A:Cross-references: GB:AE001444; GB:AE001439; NID:g4154549; PIDN:AA005635.1; PID:g415455
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0051
C:Superfamily: Helicobacter pylori hypothetical protein jhp0051

Query Match      16.9%; Score 63.5; DB 2; Length 120;
Best Local Similarity 30.6%; Pred. No.3.7;
Matches 22; Conservative 9; Mismatches 28; Indels 13; Gaps 3;

Oy      2  ESKDLSDAEPDAGEATA-----YASKDLEPOLRSVSVDELMTVLVPEY---WKWY 50
      ::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      39  QASQELGKRSPKSGKLAELFDLFEYIKDSKFERLKPSAAYDSCCKLYPQNNSQKNR 98
      ||::||::||::||::||::||::||::||::||::||::||::||::||
Oy      51  KCQLRKGQWQH 62
      :||::||::||::||::||::||::||::||::||::||::||::||::||
Db      99  RVLNR--GKYNH 108
      ||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 11
G82827
6-phosphofructokinase xF0274 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: G82827
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: G82827
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-427 <SIM>
A:Cross-references: GB:AE003880; GB:AE003849; NID:g9105080; PIDN:AAF83087.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arrida, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carter, H
as-Neto, E.; Docena, C.; El-Dorfi, H.; Facinca, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemp, E.L.; Kitajima, J.P.; Krieger, J.E.; Kurame, E.E.; Laig
chad, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Savasak

```

```

A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0274

Query Match      16.9%; Score 63.5; DB 2; Length 427;
Best Local Similarity 25.9%; Pred. No.15;
Matches 15; Conservative 13; Mismatches 25; Indels 5; Gaps 1;

Oy      2  ESKDLSDAEPDAGEATAYASKDLEPOLRSVSVDELMTVLVPEYWKWKCOLRKGW 59
      ::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      151  DNDLMTDASPFGSAAKTAVSICCALDVAAAMETSTKVE-----IYEMGRAGW 203
      ||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 12
T29030
hypothetical protein F53G12.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T29030
R:Wu, X.; Graves, T.
submitted to the EMBL Data Library, May 1997
A:Description: The sequence of C. elegans cosmid F53G12.
A:Reference number: Z20555
A:Accession: T29030
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-493 <MWX>
A:Cross-references: EMBL:AF003139; PIDN:AA54162.1; GSPDB:GN00019; CESP:F53G12.6
A:Experimental source: strain Bristol N2; clone F53G12
C:Genetics:
A:Gene: CESP:F53G12.6
A:Map position: 1
A:Insertions: 15/1; 148/3; 258/3; 338/3; 384/2; 458/3
C:Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH

Query Match      16.9%; Score 63.5; DB 2; Length 493;
Best Local Similarity 32.1%; Pred. No.18;
Matches 18; Conservative 8; Mismatches 17; Indels 13; Gaps 2;

Oy      20  YASKDLEPOL-----RSVSVDELMTVLVPEYWKWKCOLRKGWQHNRQANLNSR 71
      ||::||::||::||::||::||::||::||::||::||::||::||::||
Db      157  YAFKSCIELAHYKRNKRPYEGMTLI-----CGLARHGWLNNEDVTLNKK 203
      ||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 13
T14772
hypothetical protein DKFZp5661174.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14772
R:Bloeker, H.; Boecher, M.; Brandt, P.; Mews, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, August 1999
A:Reference number: Z18182
A:Accession: T14772
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-246 <BLD>
A:Cross-references: EMBL:AL110241
A:Experimental source: fetal kidney; clone DKFZp5661174
C:Genetics:
A>Note: DKFZp5661174.1

Query Match      16.8%; Score 63; DB 2; Length 246;
Best Local Similarity 24.4%; Pred. No.9.4;
Matches 20; Conservative 18; Mismatches 26; Indels 18; Gaps 4;

Oy      5  LDLSDAEPDAGEATAYASKD-----LEPOLRSVSVDELMTVLVPEYWKWY----- 50

```

DB 138 IDYDFVFKESKESQNSYSQSESIATCTHLOQLRTIEDIDE--NOLKDEFFKLQISIMG 195
QY 51 -KCQLRKGGQWHRQAN-LNS 70
DB 196 NKCDLSLSGGESSQNTNVLNS 217

RESULT 14

T46249
hypothetical protein DKFzp761E1312.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 02-Sep-2000
C:Accession: T46249
R:Ansorge, W.; Winkler, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23028
A:Accession: T46249
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-366 <AAA>
A:Cross-references: EMBL:AL137348
A:Experimental source: adult amygdala; clone DKFzp761E1312
C:Genetics:
A>Note: DKFzp761E1312.1
C:Superfamily: herpesvirus ribonucleoside-diphosphate reductase small chain

Query Match 16.8%; Score 63; DB 2; Length 366;
Best Local Similarity 29.4%; Pred. No. 15;
Matches 20; Conservative 14; Mismatches 28; Indels 6; Gaps 4;

QY 3 SGLDLSDAEPDAGEATATASKLEQRLSVSSVDELMTVLYPEYKMKYKCOLRKGGQW 61
DB 25 AGLDQDERSSSDTNESEIKSNE--EPLLRKSSRFVIFPIQYPDIMKMYK-QAQAQSF-- 79
QY 62 NREQANILN 69
DB 80 TAEFVVL 87

RESULT 15

T42649
hypothetical protein DKFzp434C0515.1 - human
C:Species: Homo sapiens (man)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C:Accession: T42649
R:Blöcker, H.; Boecher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, November 1999
A:Reference number: Z22230
A:Accession: T42649
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-614 <AAA>
A:Cross-references: EMBL:AL133046
A:Experimental source: adult testis; clone DKFzp434C0515
C:Genetics:
A>Note: DKFzp434C0515.1

Query Match 16.7%; Score 62.5; DB 2; Length 614;
Best Local Similarity 31.5%; Pred. No. 30;
Matches 17; Conservative 10; Mismatches 22; Indels 5; Gaps 1;
QY 3 SGLDLSDAEPDAGEATATASKLEQRLSVSSVDELMTVLYPEYKMKYKCOLRK 56
DB 307 TGVVISDVEDEAGELSRQMDL-----SCTFVREQLTNSIRKQWRIILKSHVEK 355

Search completed: October 17, 2001, 14:48:41
Job time: 287 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:51:45 ; Search time 42.92 Seconds

(without alignments)
56.667 Million cell updates/sec

Title: US-09-427-657-2_COPY_32_102

Perfect score: 375
Sequence: 1 FESGLDSDAEPPAGEATAY.....COLRKGWQHNRQANLSNR 71

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	375	100.0	419	1	VEGC_HUMAN
2	276	73.6	415	1	VEGC_MOUSE
3	64	17.1	327	1	GC4_HUMAN
4	63.5	16.9	347	1	FOS_CYPCA
5	62.5	16.7	1040	1	BOI2_YEAST
6	62	16.5	437	1	ERH_XENLA
7	61.5	16.4	441	1	PHPA_PLACH
8	61	16.3	997	1	YPR2_CAEEL
9	60	16.0	326	1	GC2_HUMAN
10	59	15.7	436	1	ERIG_XENLA
11	59	15.7	330	1	GCL_HUMAN
12	59	15.7	1005	1	BGAL_ACTPL
13	58.5	15.6	491	1	MDM2_HUMAN
14	58	15.5	363	1	YGEW_ECOLI
15	58	15.5	1097	1	LIFR_HUMAN
16	57.5	15.3	647	1	MPF1_FLAOK
17	57.5	15.3	1325	1	Y309_MYCPN
18	57	15.2	437	1	ERIG_HUMAN
19	57	15.2	517	1	ERIG_RABIT
20	57	15.2	519	1	YB3C_SCHPO
21	56.5	15.1	389	1	ARRC_RANCA
22	56.5	15.1	466	1	MDM2_MESAV
23	56.5	15.1	487	1	MDM2_CANFA
24	56.5	15.1	491	1	MDM2_HORSE
25	56.5	15.1	1224	1	Y639_YEAST
26	56.5	15.1	1224	1	Y639_YEAST
27	56.5	15.1	4644	1	DYHC_MOUSE
28	56.5	15.1	4644	1	DYHC_MOUSE
29	56	14.9	479	1	ALIN_ALICE
30	56	14.9	487	1	SYE2_THEMA
31	56	14.9	544	1	M757_ECOLI
32	55.5	14.8	256	1	MTDH_UROFA
33	55.5	14.8	424	1	CYAA_STIAU

34	55.5	14.8	623	1	DRTS_PLAVI
35	55.5	14.8	955	1	VP2_PRTV17
36	55	14.7	121	1	V121_ASFL5
37	55	14.7	329	1	YMB1_CAEEL
38	55	14.7	486	1	ALIN_ALISA
39	55	14.7	1742	1	GUWA_CALSA
40	55	14.7	1947	1	MYSC_CAEEL
41	54.5	14.5	134	1	Y322_HAEIN
42	54.5	14.5	173	1	ATP7_KLUTLA
43	54.5	14.5	389	1	ARRC_RANP1
44	54.5	14.5	485	1	HUNB_CIOAL
45	54.5	14.5	522	1	NIRK_AZOVI

ALIGNMENTS

```

RESULT 1
VEGC_HUMAN          STANDARD:      PRT:      419 AA.
AC P49767:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR (VEGF-C) (VASCULAR
DE ENDOTHELIAL GROWTH FACTOR RELATED PROTEIN) (VRF) (FLT4 LIGAND) (FLT4-
DE L).
GN VEGFC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 103-120.
RX MEDLINE=96178224; PubMed=8617204;
RA Joukov V., Pajusola K., Kaipainen A., Chliov D., Lahtinen I., Kuk E.,
RA Saksela O., Kalkkinen N., Alitalo K.;
RT "A novel vascular endothelial growth factor, VEGF-C, is a ligand for
RT the FLT4 (VEGFR-3) and KDR (VEGFR-2) receptor tyrosine kinases.";
RL EMBO J. 15:290-298(1996).
RN [2]
RP ERRATUM.
RX MEDLINE=96203094; PubMed=8612600;
RA Joukov V., Pajusola K., Kaipainen A., Chliov D., Lahtinen I., Kuk E.,
RA Saksela O., Kalkkinen N., Alitalo K.;
RL EMBO J. 15:1751-1751(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96312526; PubMed=8700872;
RA Lee J., Gray A., Yuan J., Luo S.-M., Avraham H., Wood W.I.;
RT "Vascular endothelial growth factor-related protein: a ligand and
RT specific activator of the tyrosine kinase receptor Flt4.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:1988-1992(1996).
RN [4]
RP SEQUENCE FROM N.A.
RA Fitz L., Morris J.C., Towler P.S., Long A.J., Greco R.,
RA Burgess P., Giannotti J., Charette A., Hennessey D., Kovacic S.,
RA Fitzgerald M., Scallietto H., Welch N., Neben S., Finnelly H.,
RA Zollner R., Wang J., Nickbarg E., Gassaway R., Turner K.,
RA Wood C.R.;
RT Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.
RN [5]
RP FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
RP CELL GROWTH.
RN [6]
RP SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
RN [7]
RP PTM: PROBABLY PROTEOLITICALLY PROCESSED IN THE C-TERMINUS.
RN [8]
RP SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
RN [9]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RP between the Swiss Institute of Bioinformatics and the EMBL outstation -
RP the European Bioinformatics Institute. There are no restrictions on its
RP use by non-profit institutions as long as its content is in no way
RP modified and this statement is not removed. Usage by and for commercial
RP entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@sb-sib.ch).
CC -----
CC EMBL: X94216; CA63907.1;
CC EMBL: U43142; AAA85214.1;
CC EMBL: U58112; AAB02909.1;
CC HSSP: P15692; 1YPE.
CC MIM: 601528;
CC InterPro: IPR000072;
CC InterPro: IPR002400;
CC Pfam: PF00341; PDGF_1.
CC PRINTS: PR00438; GFCYSKNOT.
CC PROSITE: PS00249; PDGF_1; 1.
CC PROSITE: PS50278; PDGF_2; 1.
CC Mitogen: Growth factor; Glycoprotein; Signal; Repeat.
CC SIGNAL: 1
CC PROPEP ? 102 POTENTIAL.
CC CHAIN 103 419 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
CC DOMAIN 275 365 4 X 24 AA TANDEM REPEATS.
CC REPEAT 275 298 1.
CC REPEAT 299 322 2.
CC REPEAT 323 346 3.
CC REPEAT 347 365 4 (PARTIAL).
CC CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 419 AA; 46883 MW; 9F598719DB3E014F CRC64;

Query Match 100.0%; Score 375; DB 1; Length 419;
Best Local Similarity 100.0%; Pred. No. 6,4e-36;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FESGIDLSDAEPDAGEATAYASKDLEQLRSVSYDELMTVLYPEYWKYKCOLRKGGMQ 60
DB 32 FESGIDLSDAEPDAGEATAYASKDLEQLRSVSYDELMTVLYPEYWKYKCOLRKGGMQ 91
OY 61 HNREQANINSR 71
DB 92 HNREQANINSR 102

RESULT 2
VEGC_MOUSE STANDARD: PRT; 415 AA.
AC P97953;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR (VEGF-C) (FLT4 LIGAND)
DE (FLT4-L).
GN VEGFC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=97164697; Pubmed=9012504;
RA Kikk E., Lybounessaki A., Taira S., Kaipainen A., Jeltsch M.,
RA Joukov V., Aitalo K.;
RT "VEGF-C receptor binding and pattern of expression with VEGFR-3
RT suggests a role in lymphatic vascular development.";
RL Development 122:3829-3837(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=97388482; Pubmed=9247316;
RA Fitz L.J., Morris J.C., Towler P., Long A., Burgess P., Greco R.,
RA Wang J., Gassaway R., Nickbarg E., Kovacic S., Charletta A.,
RA Gnanoult J., Flannerty H., Zollner R., Beier D.R., Leak L.V.,
RA Turner K.J., Wood C.R.;
RT "Characterization of murine Flt4 ligand/VEGF-C.";
RN [2]

```

```

RL Oncogene 15:613-618(1997).
CC -1- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
CC CELL GROWTH.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
CC EMBL: U73620; AAC52984.1;
CC EMBL: U58112; AAB46707.1;
CC HSSP: P15692; 1YPE.
CC MGD: MGI:109124; VEGfc.
CC InterPro: IPR000072;
CC InterPro: IPR002400;
CC Pfam: PF00341; PDGF_1.
CC PRINTS: PR00438; GFCYSKNOT.
CC PROSITE: PS00249; PDGF_1; 1.
CC PROSITE: PS50278; PDGF_2; 1.
CC Mitogen: Growth factor; Glycoprotein; Signal; Repeat.
CC SIGNAL: 1
CC PROPEP ? 98 POTENTIAL.
CC CHAIN 99 415 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
CC DOMAIN 271 361 4 X 24 AA TANDEM REPEATS.
CC REPEAT 271 294 1.
CC REPEAT 295 318 2.
CC REPEAT 319 342 3.
CC REPEAT 343 361 4 (PARTIAL).
CC CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 415 AA; 46471 MW; D9D3DD3CEC659D6 CRC64;

Query Match 73.6%; Score 276; DB 1; Length 415;
Best Local Similarity 74.6%; Pred. No. 1.7e-24;
Matches 53; Conservative 5; Mismatches 9; Indels 4; Gaps 1;

OY 1 FESGIDLSDAEPDAGEATAYASKDLEQLRSVSYDELMTVLYPEYWKYKCOLRKGGMQ 60
DB 32 FESGIDLSDAEPDAGEATAYASKDLEQLRSVSYDELMTVLYPEYWKYKCOLRKGGMQ 91
OY 61 HNREQANINSR 71
DB 92 ---OPTLNTR 98

RESULT 3
GC4_HUMAN STANDARD: PRT; 327 AA.
AC P01861;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-4 CHAIN C REGION.
GN IGHG4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=83157104; Pubmed=6299662;
RA Ellison J.W., Buxbaum J.N., Hood L.E.;
RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene.";
RL DNA 1:11-18(1981).
RN [2]
RP SEQUENCE OF 1-30 AND 81-326.

```

MEDLINE=70207560; PubMed=4192699;
 RA Pink J.R.L., Buttery S.H., de Vries G.M., Milstien C.;
 FT "Human immunoglobulin subclasses. Partial amino acid sequence of the
 RT constant region of a gamma 4 chain."
 RL Blochem. J. 117:33-47(1970).

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 CC EMBL: K01316; AAB59394.1; ALT_INIT.
 DR PIR: A02150; G4HD.
 DR MIM: 147130; -.
 DR InterPro: IPR000495; -.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 1g; 3.
 DR PROSITE: PS00290; IG_MHC; 2.
 KW Immunoglobulin domain; Immunoglobulin C region.
 FT NON_TER 1 1
 FT DOMAIN 1 98 CH1.
 FT DOMAIN 99 110 HINGE.
 FT DOMAIN 111 220 CH2.
 FT DOMAIN 221 327 CH3.
 FT DISULFD 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFD 27 83
 FT DISULFD 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFD 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFD 141 201
 FT DISULFD 247 305
 SQ SEQUENCE 327 AA; 35940 MW; 3EDB811EF208E7A CRC64;

Query Match 17.1%; Score 64; DB 1; Length 327;
 Best Local Similarity 29.7%; Pred. No. 4.1;
 Matches 19; Conservative 11; Mismatches 24; Indels 10; Gaps 2;

OY 5 LDLSAEP-----DAGEATAVASKDELEQLRSVSYDELTVLYPEYK-WKMYKCOL 54
 DB 144 VDVSODEVDFVGMVYDGVENAKTKPREOFNSTYRVSVLTVLHODWLNKREYKCKV 203
 OY 55 RKGG 58
 DB 204 SNKG 207

RESULT 4
 FOS_CYPCA STANDARD: PRT; 347 AA.
 AC P79702;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE P55-C-FOS PROTO-ONCOGENE PROTEIN (CELLULAR ONCOGENE C-FOS).
 GN CYP.
 OS *Cyprinus carpio* (Common carp).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
 CC Cyprinidae; Cyprininae; Cyprinus.
 CC NCBI_TaxID=7962;
 RN [1]
 RA SEQUENCE FROM N.A.
 RA Chang M.S., Huang C.J.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: NUCLEAR PHOSPHOPROTEIN WHICH FORMS A TIGHT BUT NON-
 CC COVALENTLY LINKED COMPLEX WITH THE C-JUN/AP-1 TRANSCRIPTION
 CC FACTOR. C-FOS HAS A CRITICAL FUNCTION IN REGULATING THE
 CC DEVELOPMENT OF CELLS DESTINED TO FORM AND MAINTAIN THE SKELETON.
 CC IT IS THOUGHT TO HAVE AN IMPORTANT ROLE IN SIGNAL TRANSDUCTION,
 CC CELL PROLIFERATION AND DIFFERENTIATION (BY SIMILARITY).

CC -!- SUBUNIT: HETERODIMER (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.
 CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. FOS SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 CC EMBL: U81505; AAB39938.1; -.
 DR HSSP: P01100; IPOS.
 DR InterPro: IPR000837; -.
 DR InterPro: IPR001871; -.
 DR Pfam: PF00170; BZIP; 1.
 DR PRINTS: PR00042; LEUZIPPROF.
 DR PROSITE: PS00036; BZIP_BASIC; 1.
 KW Proto-oncogene; Nuclear protein; Phosphorylation; DNA-binding.
 FT DNA_BIND 115 136 BASIC MOTIF.
 FT DOMAIN 141 169 LEUCINE-ZIPPER.
 SQ SEQUENCE 347 AA; 37586 MW; A8F8781044A8CE1 CRC64;

Query Match 16.9%; Score 63.5; DB 1; Length 347;
 Best Local Similarity 37.5%; Pred. No. 5;
 Matches 18; Conservative 7; Mismatches 22; Indels 1; Gaps 1;
 OY 1 FESGDLSDPEADAGEATVASKDLEQLRSVSYDELTVLYPEYK 48
 DB 231 FGSYVEISDLEPTLESLELAKAELEFRASVPDWD-ISSSIYADWE 277

RESULT 5
 BOI2_YEAST STANDARD: PRT; 1040 AA.
 AC P39969;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE BOI2 PROTEIN (BEB1 PROTEIN).
 GN BOI2 OR BEB1 OR YER114C.
 OS *Saccharomyces cerevisiae* (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 CC NCBI_TaxID=4932;
 RN [1]
 RA SEQUENCE FROM N.A.
 RA Matsui Y., Matsui R., Toh E. A.;
 RL Submitted (XXX-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RA SEQUENCE FROM N.A.
 RA STRAIN=S288C / AB972;
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
 RA Aviles E., Beno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
 RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
 RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
 RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
 RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
 RA Taylor P., Wei Y., Yelton M., Borstein D., Davis R.W.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: BINDS TO THE BEB1 PROTEIN. INVOLVED IN BUD FORMATION.
 CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 SAM DOMAIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

Query Match 16.7% Score 62.5; DB 1, length 1040;
Best Local Similarity 32.7%; Pred No. 23;
Matches 16; Conservative 11; Mismatches 15; Indels 7; Gaps 3

OY 4 GLDLS-SPAEPDGAETAYASKDLDEQLRSVSDVELMTLYP---EYW 47
||| :
||| :
||| :

zb 220 GLDVNSESKEPTTSSSTGFLNDGLNDOALINGID--TRKLMPVEAEFW 266
||||| :
||||| :

RESULT	6			
ID	EF1H_XENLA	STANDARD:	PRT:	437 AA.
AC	091375;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	15-JUL-1998 (Rel. 36, Last annotation update)			
DE	ELONGATION FACTOR 1-GAMMA TYPE 2 (EF-1-GAMMA) (P47).			
OS	Xenopus laevis (African clawed frog).			
OC	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;			
OC	Xenopodinae; Xenopus.			
OX	NCBI_TaxID=8355;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Oocyte;			
RX	MEDLINE=94155465; PubMed=811972;			
RA	Morales J., Bassez T., Cormier P., Mulner-Lorillon O., Belle R.,			
RA	Osborne H.B.,			
RT	"Expression of elongation factor 1 alpha (EF-1 alpha) and 1 beta			
RT	gamma (EF-1 beta gamma) are uncoupled in early Xenopus embryos";			
RL	Dev. Genet. 14:440-448(1993).			
CC	-1- FUNCTION: PROBABLY PLAYS A ROLE IN ANCHORING THE COMPLEX TO OTHER			
CC	CELLULAR COMPONENTS.			
CC	-1- SUBUNIT: EF-1 IS COMPOSED OF FOUR SUBUNITS: ALPHA, BETA,			
CC	DELTA, AND GAMMA.			
CC	-1- SIMILARITY: TO OTHER SPECIES' GAMMA SUBUNIT AS WELL AS TO THE			
CC	N-TERMINAL OF SOME EUKARYOTIC AMINO ACID-TRNA SYNTHETASES.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	at the European Bioinformatics Institute. There are no restrictions on			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce)			
CC	or send an email to license@isb-sib.ch .			
CC	-----			
DR	EMBL; S69726; AAB29958.1; -;			
DR	InterPro; IPR000521; -;			
DR	InterPro; IPR001662; -;			
DR	Pfam; PF00647; EFIG_domain; 1.			
DR	Pfam; PF00043; GST; 1.			
DR	PROSITE; PS50040; EFIG; 1.			
FW	Elongation factor; Protein biosynthesis.			

Query Match	16.5%	Score 62;	DB 1;	Length 437;
Best Local Similarity	32.3%	Pred. No. 9.6;		
Matches	21;	Conservative	11;	Mismatches 21;
			Indels	12;
			Gaps	4

```

OY  2  EGGDLSQ----AEPDAGEATVASKD---LEEDLRSSVSDVELMTVLYPEYKMYKCOL  54
      | | | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db  260 EDDDLDESKALAAEPKSKNDPYAHLPKSSFINDPKRRKYSNBDTL-TYALPYFWMEHFE---  315
      | | | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY  55 RKKGW  59
      | | | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db  316 -KEGW  319

```

RESULT	7
PHPA_PLACH	
ID	PHPA_PLACH STANDARD; PRT; 441 AA.
AC	Q02752;
DT	01-JUL-1993 (Rel. 26, Created)
DT	01-JUL-1993 (Rel. 26, last sequence update)
DT	01-JUN-1994 (Rel. 29, last annotation update)
DE	ACIDIC PHOSPHOPROTEIN PRECURSOR (50 KDA ANTIGEN).
GN	PCPMAL
OS	Plasmodium chabaudi.
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_taxid	5825;
TX	

RP	SEQUENCE FROM N.A.
RC	STRAIN-IP-PC1
RX	MEDLINE=93116806; PubMed=1475002;
RA	Deleersnijder W., Prasomstitt P., Tungradubkul S., Hendrix D.,
RA	Hamers-Casterman C., Hamers R.;
RT	"Structure of a Plasmodium chabaudi acidic phosphoprotein that is
RT	associated with the host erythrocyte membrane.";
RL	Mol. Biochem. Parasitol. 56:59-68(1992).
CC	-1- FUNCTION: DURING INFECTION, THIS PHOSPHOPROTEIN PROBABLY MODULATES
CC	THE STRUCTURE OF THE RED CELL MEMBRANE TO THE ADVANTAGE OF THE
CC	PARASITE, ALTHOUGH ITS PRECISE FUNCTION IS NOT KNOWN.
CC	-1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN ON THE
CC	CYTOPLASMIC FACE OF THE HOST ERYTHROCYTE MEMBRANE.
CC	-1- MISCELLANEOUS: ASSOCIATED WITH THE HOST RED CELL MEMBRANE
CC	THROUGHOUT THE ENTIRE ERYTHROCYTIC CYCLE.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-slb.ch/announce/).
CC	or send an email to license@isb-slb.ch).
CC	-----
DR	EMBL: M95789; AAA29732.1; -.
DR	PIR: A48455; A48455
DR	phosphorylation; Signal; Antigen; Membrane; Repeat; Erythrocyte.
KW	SIGNAL
FT	1 15
FT	16 441
FT	186 313
FT	186 193
FT	194 201
FT	202 209
FT	210 217
FT	218 217
FT	218 225
FT	226 233
FT	234 241
FT	242 249
FT	250 257
FT	258 265
FT	266 273
FT	274 281
FT	282 289
FT	290 297
FT	1 15
FT	OR 24 (POTENTIAL)
FT	ACIDIC PHOSPHOPROTEIN.
FT	16 X 8 AA TANDEM REPEATS.
FT	1-1.
FT	1-2.
FT	1-3.
FT	1-4.
FT	1-5.
FT	1-6.
FT	1-7.
FT	1-8.
FT	1-9.
FT	1-10.
FT	1-11.
FT	1-12.
FT	1-13.
FT	1-14.

DR MM, 147110; -
 DR InterPro; IPR000495; -
 DR InterPro; IPR003006; -
 DR Pfam; PF00047; 1g; 3.
 DR PROSITE; PS00290; IG_MHC; 2.
 KW Immunoglobulin domain; Immunoglobulin C region.
 FT NON_TER 1 1
 FT DOMAIN 1 98 CH1.
 FT DOMAIN 99 110 HINGE.
 FT DOMAIN 111 219 CH2.
 FT DOMAIN 220 326 CH3.
 FT DISULFD 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFD 27 83 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFD 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFD 103 103 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFD 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFD 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFD 140 200 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFD 246 304
 FT SITE 156 156
 FT MOD_RES 326 326 AT OR NEAR THE COMPLEMENT-BINDING SITE.
 FT VARIANT 60 60 REMOVED POST-TRANSLATIONALLY (PROBABLY).
 FT S -> A (IN MYELOMA PROTEINS TIL & ZIE).
 FT /FTID=VAR.003889
 SQ SEQUENCE 326 AA; 35884 MW; 8310878C6878CF9C CRC64;

Query Match 16.0%; Score 60; DB 1; Length 326;
 Best Local Similarity 28.1%; Pred. No. 12;
 Matches 18; Conservative 12; Mismatches 24; Indels 10; Gaps 2;

QY 5 LDLSAEP-----DAGEATATVSKDLEQLRSVSVDELMTVLYPEYKWKYKCOL 54
 DB 143 VDVSHEDEVOFNMVYDGVEMHNAKTPREQFNSTFRVSVLTIVHODWLNGKEYCKV 202
 QY 55 RKGG 58
 DB 203 SNKG 206

RESULT 10
 EF1G_XENLA STANDARD; PRT; 436 AA.
 AC P26642; O91374;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE ELONGATION FACTOR 1-GAMMA TYPE 1 (EF-1-GAMMA) (P47).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 OC Xeropodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=92093634; PubMed=1754404;
 RA Cormier P., Osborne H.B., Morales J., Bassez T., Pouthe R.,
 RA Mazabraud A., Mulner-Lorillon O., Belle R.;
 RT "Molecular cloning of Xenopus elongation factor 1 gamma, major
 RT M-phase promoting factor substrate";
 RT Nucleic Acids Res. 19:6644-6644(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oocyte;
 RX MEDLINE=94155465; PubMed=8111972;
 RA Morales J., Bassez T., Cormier P., Mulner-Lorillon O., Belle R.,
 RA Osborne H.B.;
 RT "Expression of elongation factor 1 alpha (EF-1 alpha) and 1 beta
 RT gamma (EF-1 beta gamma) are uncoupled in early Xenopus embryos.";
 RT Dev. Genet. 14:440-448(1993).
 CC -1- FUNCTION: PROBABLY PLAYS A ROLE IN ANCHORING THE COMPLEX TO OTHER
 CC CELLULAR COMPONENTS.
 CC -1- SUBUNIT: EF-1 IS COMPOSED OF FOUR SUBUNITS: ALPHA, BETA,

CC DELTA, AND GAMMA.
 CC -1- SIMILARITY: TO OTHER SPECIES' GAMMA SUBUNIT AS WELL AS TO THE
 CC N-TERMINAL OF SOME EUKARYOTIC AMINO ACID-TRNA SYNTHETASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X62508; CAA44367.1; -
 DR EMBL; S69724; AAB29957.1; -
 DR PIR; S18051; S18051.
 DR PIR; S20060; S20060.
 DR InterPro; IPR000521; -
 DR InterPro; IPR001662; -
 DR Pfam; PF00647; EF1G-domain; 1.
 DR Pfam; PF00043; GST; 1.
 DR PROSITE; PS00040; EF1G; 1.
 KW Elongation factor; Protein biosynthesis.
 FT CONFLICT 134 134 G -> E (IN REF. 2).
 FT SEQUENCE 436 AA; 49791 MW; 8785C1E80578B131 CRC64;

Query Match 16.0%; Score 60; DB 1; Length 436;
 Best Local Similarity 32.3%; Pred. No. 16;
 Matches 21; Conservative 10; Mismatches 22; Indels 12; Gaps 4;

QY 2 ESDLSLD---AEPDAGEATATVASKD---LEQLRSVSVDELMTVLYPEYKWKYKCOL 54
 DB 259 EDDLSESKALAAEPKSKDPYAHLPKSSFMDEPKRYSNEDTL-TVALPYWEHFD--- 314
 QY 55 RKGGW 59
 DB 315 -KEGW 318

RESULT 11
 GCL_HUMAN STANDARD; PRT; 330 AA.
 AC P01857;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG GAMMA-1 CHAIN C REGION.
 GN IGHL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=82274238; PubMed=6287432;
 RA Ellison J.W., Berson B.J., Hood L.E.;
 RT "The nucleotide sequence of a human immunoglobulin C gamma1 gene";
 RT Nucleic Acids Res. 10:4071-4079(1982).
 RN [2]
 RP SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).
 RX MEDLINE=71064024; PubMed=5489771;
 RA Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,
 RA Wexdal M.J., Edelman G.M.;
 RT "The covalent structure of a human gamma G-immunoglobulin. VII. Amino
 RT acid sequence of heavy-chain cyanogen bromide fragments H1-H4";
 RT Biochemistry 9:3161-3170(1970).
 RN [3]
 RP SEQUENCE OF 136-329 (EU).
 RX MEDLINE=71064025; PubMed=5530842;
 RA Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,
 RA Edelman G.M.;
 RT "The covalent structure of a human gamma G-immunoglobulin. 8. Amino
 RT acid sequence of heavy-chain cyanogen bromide fragments H5-H7";

RL Biochemistry 9:3171-3181(1970).
 RN [4]
 RN SEQUENCE (MYELOMA PROTEIN NIE).
 RX MEDLINE=77070269; PubMed=826475;
 RA Ponstingl H., Hilschmann N.;
 RT "The rule of antibody structure. The primary structure of a
 RT monoclonal IgG1 immunoglobulin (myeloma protein Nie). III. The
 RT chymotryptic peptides of the H-chain, alignment of the tryptic
 RT peptides and discussion of the complete structure.";
 RL Hoppe-seyler's Z. Physiol. Chem. 357:1571-1604(1976).
 RN [5]
 RN SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.
 RX MEDLINE=83289131; PubMed=6884994;
 RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;
 RT "Three-dimensional structure determination of antibodies. Primary
 RT structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";
 RL Hoppe-seyler's Z. Physiol. Chem. 364:713-747(1983).
 RN [6]
 RN DISULFIDE BONDS.
 RX MEDLINE=71064027; PubMed=4923144;
 RA Gail W.E., Edelman G.M.;
 RT "The covalent structure of a human gamma G-immunoglobulin. X.
 RT Intrachain disulfide bonds.";
 RL Biochemistry 9:3188-3196(1970).
 RN [7]
 RN DISULFIDE BONDS.
 RX MEDLINE=77070267; PubMed=1002129;
 RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;
 RT "Rule of antibody structure. The primary structure of a monoclonal
 RT IgG1 immunoglobulin (myeloma protein Nie). I: Purification and
 RT characterization of the protein, the L- and H-chains, the
 RT cyanogen bromide cleavage products, and the disulfide bridges.";
 RL Hoppe-seyler's Z. Physiol. Chem. 357:1515-1540(1976).
 RN [8]
 RN X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
 RX MEDLINE=81208100; PubMed=7236608;
 RA Delsenhofer J.;
 RT "Crystallographic refinement and atomic models of a human Fc fragment
 RT and its complex with fragment B of protein A from Staphylococcus
 RT aureus at 2.9- and 2.8-A resolution.";
 RL Biochemistry 20:2361-2370(1981).
 CC -I- MISCELLANEOUS: NIE HAS THE G1M(17) ALLOTYPIC MARKER, 97-K, & THE
 CC G1M(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE G1M(3)
 CC MARKER & THE G1M (NON-1) MARKERS.
 CC -I- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF
 CC 35,116,198,269 & 272.
 CC -I- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES
 CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES
 CC 268-272.
 CC -I- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF
 CC RESIDUES 198,267&272.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: J00228; AAC82527.1; ALT_INIT.
 DR PIR: A02146; GHU.
 DR PDB: 1FC1; 15-JUL-92.
 DR PDB: 1FC2; 15-JUL-92.
 DR MIM: 147100;
 DR InterPro: IPR000495;
 DR InterPro: IPR003006;
 DR Pfam: PF00047; 1g; 3.
 DR PROSITE: PS00290; IG_MHC; 2.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
 KW 3D-structure.
 FT NON_TER 1 1
 FT DOMAIN 1 98 CH1.

FT DOMAIN 99 110 HINGE.
 FT DOMAIN 111 223 CH2.
 FT DOMAIN 224 330 CH3.
 FT DISULFID 27 83
 FT DISULFID 103 103 INTERCHAIN (WITH LIGHT CHAIN).
 FT DISULFID 109 109 INTERCHAIN (WITH HEAVY CHAIN).
 FT DISULFID 112 112 INTERCHAIN (WITH HEAVY CHAIN).
 FT DISULFID 144 204
 FT DISULFID 250 308
 FT CARBOHYD 180 180
 FT MOD_RES 350 330
 FT VARIANT 97 97
 FT VARIANT 239 239
 FT VARIANT 241 241
 FT STRAND 123 126
 FT HELIX 130 134
 FT TURN 136 137
 FT STRAND 141 148
 FT TURN 158 162
 FT TURN 163 164
 FT STRAND 165 166
 FT STRAND 175 178
 FT STRAND 183 190
 FT HELIX 193 197
 FT TURN 198 199
 FT STRAND 202 206
 FT STRAND 215 219
 FT STRAND 227 227
 FT STRAND 230 234
 FT HELIX 238 240
 FT TURN 241 242
 FT STRAND 245 256
 FT STRAND 260 266
 FT TURN 267 268
 FT STRAND 269 270
 FT STRAND 274 276
 FT STRAND 280 281
 FT TURN 283 284
 FT STRAND 287 296
 FT HELIX 297 301
 FT TURN 302 303
 FT STRAND 306 312
 FT TURN 313 314
 FT TURN 316 317
 FT STRAND 320 324
 SO SEQUENCE 330 AA; 36106 MW; 3770EE106C2FA3D CRC64;
 Query Match 15.7%; Score 59; DB 1; Length 330;
 Best Local Similarity 30.0%; Pred. No. 16;
 Matches 18; Conservative 11; Mismatches 21; Indels 10; Gaps 2;
 QY 5 IDLSAEP-----DGEATAYASKDLEQLRSVSYDEMTLYPY--WKWYKOL 54
 DB 147 VDVSHDEPVKFNWYVDGVEVHNAKTKPREQYNSYRVSVLYVLDHDLNGKERYCKV 206
 RESULT 12
 BGAL_ACTPL STANDARD; PRT: 1005 AA.
 ID BGAL_ACTPL
 AC P70753;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 DE BETA-GALACTOSIDASE (EC 3.2.1.23) (LACTASE).
 GN LACZ.
 OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Actinobacillus.
 OX NCBI_TaxID=715;

FN	[1]	SEQUENCE FROM N.A.
RP		
RC	STRAIN=CM5 / SEROTYPE 1;	
RX	MEDLINE=97372558; PubMed=9228778;	
RA	Anderson T.J., MacInnes J.I.;	
RT	"Expression and phylogenetic relationships of a novel lact homologue	
RL	from <i>Actinobacillus pleuropneumoniae</i> ,";	
CC	FEMS Microbiol. Lett. 152:117-123(1997).	
CC	-I- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL, NON-REDUCING BETA-D-	
CC	GALACTOSE RESIDUES IN BETA-D-GALACTOSIDES.	
CC	-I- SIMILARITY: BELONGS TO FAMILY 2 OF GLYCOSYL HYDROLASES.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (see http://www.isb-slb.ch/announce/	
CC	or send an email to license@isb-slb.ch).	
CC	-----	
DR	EMBL; U62625; AAB17954.1; -	
DR	InterPro; IPR001649; -	
DR	Pfam; PF00703; Glyco_hydro_2; 1.	
DR	PRINTS; PR00732; GLHYDRBASE2.	
DR	PROSITE; PS00719; GLYCOSYL_HYDROL_F2_1; 1.	
DR	PROSITE; PS00608; GLYCOSYL_HYDROL_F2_2; 1.	
KM	Hydrolase; Glycosidase.	
FT	ACT_SITE 455 455	
FT	ACT_SITE 526 526	
FT	PROTON DONOR (BY SIMILARITY).	
FT	NUCLEOPHILE (BY SIMILARITY).	
QO	SEQUENCE 1005 AA; 117008 MW; 12254482581C710E CRC64;	

Query Match	15.7%	Score	59	DB	1	Length	1005
Best Local	29.3%	Pred. No.	56				
Matches	12	Conservative	9	Mismatches	20	Indels	0
						Gaps	0

QY 4 GLDLSDAEPDAGEATAYASKDLEEQLRSVSSVDELMTVLYP 44
 :: :: | | :: :: | |
 Db 317 GVNRRHSDPKGTGYAITYYAQAHKDLQLMKQHNINAIRTAHYR 357

	RESULT	13
MDM2_HUMAN	ID	MDM2_HUMAN
AC	Q00967;	Q13226; Q13297; Q13298; Q13299; Q13300; Q13301; Q9UGI3;
AC	Q50MT8;	
DT	01-APR-1993 (Rel.	25, Created)
DT	01-APR-1993 (Rel.	25, Last sequence update)
DT	01-OCT-2000 (Rel.	40, Last annotation update)
DE	UBIQUITIN-PROTEIN LIGASE E3 MDM2 (EC 6.3.2.-) (P53-BINDING PROTEIN MDM2) (ONCOPROTEIN MDM2) (DOUBLE MINUTE 2 PROTEIN) (HDM2).	
GN	MDM2	
OS	Homo sapiens (human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCB1_TaxID=9606;	
RN	[1]	
RN	SEQUENCE FROM N.A.	
RP	MEDLINE=92310576; PubMed=1614537;	
RX	Oliner J.D., Kinzler K.W., Melzer P.S., George D.L., Vogelstein B., "Amplification of a gene encoding a p53-associated protein in human sarcomas."	
RT	Nature 358:80-83(1992).	
RL	[2]	
RN	SEQUENCE FROM N.A. (ISOFORMS MDM2-A; -B; -C; -D AND -E).	
RP	TISSUE=Ovarian carcinoma;	
RC	MEDLINE=96313107; PubMed=8705862;	
RX	Sigalas I., Calvert A.H., Anderson J.J., Neal D.E., Lunec J.; "Alternately spliced mdm2 transcripts with loss of p53 binding domain sequences: transforming ability and frequent detection in human cancer."	
RT	Nat. Med. 2:912-917(1996).	
RL		

[3]
SEQUENCE FROM N.A. (ISOFORM MDM2-ALPHA).
MEDLINE=20065171; PubMed=10597303;
Veldhoven N., Metcalfe S., Milner J.;
"A novel exon within the mdm2 gene modulates translation initiation in vitro and disrupts the p53-binding domain of mdm2 protein.";
Oncogene 18:7026-7033(1999).
[4]
SEQUENCE OF 6-491 FROM N.A. (ISOFORM MDM2-A1).
Liang H., Atkins H., Abdel-Fattah R., Snaeyun R., Lunec J.;
"Genomic Organisation of the Human MDM2 Oncogene and Relationship to its Alternatively Spliced mRNA's";
submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
[5]
SEQUENCE OF 1-24 FROM N.A.
MEDLINE=95380270; PubMed=7651818;
Zanderman A., Flusberg D., Haupt Y., Barak Y., Oren M.;
"A functional p53-responsive intronic promoter is contained within the human mdm2 gene";
Nucleic Acids Res. 23:2584-2592(1995).
[6]
SEQUENCE OF 1-9 FROM N.A.
MEDLINE=97413643; PubMed=9270029;
Landers J.E., Cassel S.L., George D.L.;
"Translational enhancement of mdm2 oncogene expression in human tumor cells containing a stabilized wild-type p53 protein";
Cancer Res. 57:3562-3568(1997).
[7]
SEQUENCE OF 301-481 FROM N.A.
Tabert H.W., Kappler M., Meyer A., Bartel F., Schloft T., Bache M., Schmidt H., Wuehl P.;
"A Mbo II polymorphism in exon 11 of the human MDM2 gene occurring in normal blood donors and in soft tissue sarcoma patients";
submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
[8]
MUTAGENESIS OF CYS-464.
MEDLINE=98111004; PubMed=9450543;
Honda R., Tanaka H., Yasuda H.;
"Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53.";
FEBS Lett. 420:25-27(1997).
[9]
MUTAGENESIS OF CYS-449.
MEDLINE=20190101; PubMed=10723139;
Honda R., Yasuda H.;
"Activity of MDM2, a ubiquitin ligase, toward p53 or itself is dependent on the RING finger domain of the ligase.";
Oncogene 19:1473-1476(2000).
[10]
MUTAGENESIS.
MEDLINE=20187618; PubMed=10722742;
Fang S., Jensen U.P., Ludwig R.L., Vonsden K.H., Weissman A.M.;
"Mdm2 is a chem. finger-finger-dependent ubiquitin protein ligase for itself and p53";
J. Biol. Chem. 275:8945-8951(2000).
[11]
MUTAGENESIS OF CYS-441 AND CYS-478.
MEDLINE=20076498; PubMed=10608892;
Sharp D.A., Krawtowitz S.A., Sank M.J., George D.L.;
"Stabilization of the MDM2 oncoprotein by interaction with the structurally related MDMX protein";
J. Biol. Chem. 274:38189-38196(1999).
[12]
NUCLEOLAR LOCALIZATION SIGNAL.
MEDLINE=20173879; PubMed=10707090;
Idrissi M.A.E., Ashcroft M., Kubbutat M.H.G., Vonsden K.H.;
"Identification of a cryptic nucleolar-localization signal in MDM2";
Nat. Cell Biol. 2:179-181(2000).
[13]
PHOSPHORYLATION BY ATM.
MEDLINE=20079591; PubMed=10611322;
Khosravi R., Maya R., Gottlieb T., Oren M., Shiloh Y., Shkedy D.;
"Rapid ATM-dependent phosphorylation of MDM2 precedes p53 accumulation in response to DNA damage";

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:50:54 ; Search time 128.06 Seconds
(Without alignments)
73.354 Million cell updates/sec

Title: US-09-427-657-2_COPY_32_102

Perfect score: 375
Sequence: 1 FESGILSDAEPPDAGEATAYCOLRKGGMHNEQANLSNR 71

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_16:*
1: sp._archaea:*
2: sp._bacteria:*
3: sp._fungi:*
4: sp._human:*
5: sp._invertebrate:*
6: sp._mammal:*
7: sp._mhc:*
8: sp._organelle:*
9: sp._phage:*
10: sp._plant:*
11: sp._rodent:*
12: sp._unclassified:*
13: sp._vertebrate:*
14: sp._virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	316	84.3	420	6 09XS50	09XS50 bos taurus
2	284	75.7	418	13 057352	057352 coturnix co
3	88	23.5	326	11 035251	035251 rattus norv
4	87	23.2	354	4 043915	043915 homo sapien
5	86	22.9	358	11 P97946	P97946 mus musculu
6	69	18.4	786	5 021027	021027 caenorhabd
7	67.5	18.0	443	5 09VMY5	09VMY5 drosophila
8	67	17.9	453	2 09JZY9	09JZY9 neisseria m
9	67	17.9	453	2 09JVO8	09JVO8 neisseria m
10	67	17.9	1299	10 09STEO	09STEO brassica m
11	66	17.6	602	5 09W397	09W397 drosophila
12	66	17.6	703	2 09K022	09K022 vibrio chol
13	65	17.3	1612	5 09VY02	09VY02 drosophila
14	65	17.3	2408	5 022184	022184 caenorhabd
15	64.5	17.2	4488	11 0902H1	0902H1 mus musculu
16	64	17.1	284	1 09V103	09V103 pyrococcus
17	64	17.1	926	5 09W057	09W057 drosophila
18	64	17.1	1367	10 09SFB6	09SFB6 arabidopsis
19	64	17.1	1368	10 081809	081809 arabidopsis

20	64	17.1	1368	10 09LJDB	09LJDB arabidopsis
21	64	17.1	1403	10 09SUZ1	09SUZ1 arabidopsis
22	63.5	16.9	120	2 09ZN09	09ZN09 helicobacte
23	63.5	16.9	427	2 09PCM4	09PCM4 xylella fas
24	63.5	16.9	493	5 001798	001798 caenorhabd
25	63.5	16.9	1778	5 09GUX5	09GUX5 anthraea p
26	63	16.8	246	4 09UFV5	09UFV5 homo sapien
27	63	16.8	351	4 09NFW3	09NFW3 homo sapien
28	63	16.8	351	4 09NPD6	09NPD6 homo sapien
29	63	16.8	366	4 09NTD8	09NTD8 homo sapien
30	62.5	16.7	385	11 09ER38	09ER38 mus musculu
31	62.5	16.7	614	4 09Y2D8	09Y2D8 homo sapien
32	62.5	16.7	614	4 09UX00	09UX00 homo sapien
33	62	16.5	210	2 09JY25	09JY25 neisseria m
34	62	16.5	210	2 09JY21	09JY21 neisseria m
35	62	16.5	371	2 056080	056080 salmonella
36	62	16.5	371	2 09E015	09E015 salmonella
37	62	16.5	1729	3 010287	010287 schizosacch
38	61.5	16.4	138	5 096626	096626 dictyostell
39	61.5	16.4	192	11 09QY58	09QY58 mus musculu
40	61.5	16.4	871	2 09RJ81	09RJ81 streptomyce
41	61	16.3	1220	10 024015	024015 lycopersico
42	60.5	16.1	474	2 09I189	09I189 pseudomonas
43	60	16.0	304	1 049156	049156 methanother
44	60	16.0	441	4 09H993	09H993 homo sapien
45	60	16.0	1119	5 09W224	09W224 drosophila

ALIGNMENTS

RESULT	1	PRELIMINARY	PRT	420 AA.
09XS50				
AC 09XS50				
DT 01-NOV-1999 (TEMBLrel. 12, Created)				
DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)				
DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)				
DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR.				
OS Bos taurus (Bovine).				
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OC Bovidae; Bovinae; Bos.				
OX NCBI_TaxID=9913;				
RN [1]				
RP SEQUENCE FROM N.A.				
RC TISSUE=HEART;				
RA Liu X., Yonekura H., Yamagishi S., Yamamoto Y., Yamamoto H.;				
RT "Structure and expression of bovine VEGF family.";				
RL Submitted (MAY-1987) to the EMBL/GenBank/DBJ databases.				
DR EMBL: AB004275; BAA77687.1; -.				
DR HSSP: P15692; IVP.				
DR InterPro: IPR000072; -.				
DR Pfam: PF00341; PDGF_1.				
DR PROSITE: PS00249; PDGF_1; 1.				
DR PROSITE: PS00278; PDGF_2; 1.				
DR SMART: SM00141; PDGF; 1.				
KW Signal.				
FT SIGNAL	1	20	POTENTIAL.	
FT CHAIN	21	420	VASCULAR ENDOTHELIAL GROWTH FACTOR C.	
SQ SEQUENCE	420 AA;	46681 MW;	58BA84317A3C8E2D CRC64;	

Query Match 84.3%; Score 316; DB 6; Length 420;

Best local Similarity 83.1%; Pred. No. 1.4e-27;

Matches 59; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY	1	FESGILSDAEPPDAGEATAYASKDLEQLRSVSVDELMTVLYPEYWKMKCOLRRGGMQ	60
DB	33	FESGILSDAEPPDAGEATAYASKDLEQLRSVSVDELMTVLYPEYWKMKCOLRRGGMQ	92
OY	61	HNREQANLSNR 71	

```

Db      93 HSTEQANTNIR 103

RESULT  2
ID      057352      PRELIMINARY:      PRT:      418 AA.
AC      057352;
DT      01-JUN-1998 (TREMBlrel. 06, Created)
DT      01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT      01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE      VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR.
GN      VEGF-C.
OS      Coturnix coturnix japonica (Japanese quail).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC      Coturnix.
OX      NCBI_TaxID=93934;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=98167900; PubMed=9435294;
RA      Eichmann A., Corbel C., Jaffredo T., Breant V., Joukov V., Kumar V.,
RA      Allitalo K., Le Douarin N.M.;
RT      "Avian VEGF-C: cloning, embryonic expression pattern and stimulation
RT      of the differentiation of VEGFR2-expressing endothelial cell
RT      precursors."
RT      Development 125:743-752(1998).
RL      EMBL; Y15837; CAA75799.1; -.
DR      HSSP; P15692; IYPP.
DR      InterPro: IPR000072; -.
DR      InterPro: IPR002400; -.
DR      Pfam; PF00341; PDGF_1.
DR      PRINTS; PR00438; GFCSKNOT.
DR      PRODOM; PD001629; -. 1.
DR      PROSITE; PS00249; PDGF_1; 1.
DR      PROSITE; PS50278; PDGF_2; 1.
DR      SMART; SM00141; PDGF; 1.
KW      Signal.
FT      SIGNAL.
FT      CHAIN.
SQ      SEQUENCE 418 AA; 46839 MW; 099BFCC79151BF28 CRC64;

Query Match
Best Local Similarity 75.7%; Score 284; DB 13; Length 418;
Matches 53; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

QY      1 FESGIDLSDAEPDAGEATVYASNDLEEQRLSVSSVDELTMTVLYPEYKMKYKCOLRKGMQ 60
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB      31 YESGHGYEEEPGAGEPKHAASKDLEQLRSVSVDELMTFTVLYPEYKMKYKCOLRKGMQ 90
QY      61 HNRQANLNSR 71
      |||| :|||
DB      91 HNRHSSSDPTR 101

RESULT  3
ID      035251      PRELIMINARY:      PRT:      326 AA.
AC      035251;
DT      01-JAN-1998 (TREMBlrel. 05, Created)
DT      01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT      01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE      VASCULAR ENDOTHELIAL GROWTH FACTOR D.
GN      VEGF-D.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=SPRAGUE DAWLEY;
RX      MEDLINE=97349118; PubMed=9205122;
      Yamada Y., Nezu J., Shimane M., Hirata Y.;

```

```

RT      "Molecular cloning of a novel vascular endothelial growth factor,
RT      VEGF-D."
RT      Genomics 42:483-488(1997).
DR      EMBL; AF014827; AAB66557.1; -.
DR      HSSP; P15692; IYPP.
DR      InterPro: IPR000072; -.
DR      Pfam; PF00341; PDGF; 1.
DR      PRODOM; PD001629; -. 1.
DR      PROSITE; PS00249; PDGF_1; 1.
DR      PROSITE; PS50278; PDGF_2; 1.
DR      SMART; SM00141; PDGF; 1.
SQ      SEQUENCE 326 AA; 37112 MW; 1261AFA373596C00 CRC64;

Query Match
Best Local Similarity 23.5%; Score 88; DB 11; Length 326;
Matches 16; Conservative 16; Mismatches 8; Indels 6; Gaps 1;

QY      26 EQQLRSVSVDELMTVLYPEYKMKYKCOLRKGMQHNRQANLNSR 71
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB      46 EQQIRASTLEELQVHAHSEDMKLMRCRLK-----KSLANVDSR 85

RESULT  4
ID      043915      PRELIMINARY:      PRT:      354 AA.
AC      043915;
DT      01-JUN-1998 (TREMBlrel. 06, Created)
DT      01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT      01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE      GROWTH FACTOR FIGF.
GN      FIGF OR VEGF-D.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=98140120; PubMed=9479493;
RA      Rocchigiani M., Lestingi M., Luddi A., Orlandini M., Franco B.,
RA      Rossi E., Ballabio A., Zuffardi O., Oliviero S.;
RT      "Human FIGF: cloning, gene structure, and mapping to chromosome Xp22.1
RT      between the FIGA and the GRPR genes."
RT      Genomics 47:207-216(1998).
RL      [2]
RP      SEQUENCE FROM N.A.
RC      TISSUE=LUNG;
RX      MEDLINE=97349118; PubMed=9205122;
      Yamada Y., Nezu J., Shimane M., Hirata Y.;
RT      "Molecular cloning of a novel vascular endothelial growth factor,
RT      VEGF-D."
RT      Genomics 42:483-488(1997).
RL      [3]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=98118549; PubMed=9435229;
RA      Achen M.G., Jeltsch M., Kukk E., Maekinen T., Vitali A., Wilks A.F.,
RA      Allitalo K., Stacker S.A.;
RT      "Vascular endothelial growth factor D (VEGF-D) is a ligand for the
RT      tyrosine kinases VEGF receptor 2 (Flk1) and VEGF receptor 3 (Flt4).";
RT      Proc. Natl. Acad. Sci. U.S.A. 95:548-553(1998).
DR      EMBL; Y12864; CAA73371.1; -.
DR      EMBL; Y12865; CAA73371.1; JOINED.
DR      EMBL; Y12866; CAA73371.1; JOINED.
DR      EMBL; Y12867; CAA73371.1; JOINED.
DR      EMBL; Y12868; CAA73371.1; JOINED.
DR      EMBL; Y12869; CAA73371.1; JOINED.
DR      EMBL; Y12870; CAA73371.1; JOINED.
DR      EMBL; D89630; BAA24264.1; -.
DR      EMBL; AJ000185; CAA03942.1; -.
DR      EMBL; Y12863; CAA73370.1; -.
DR      HSSP; P15692; IYPP.
DR      InterPro: IPR000072; -.
DR      Pfam; PF00341; PDGF; 1.

```


RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borikova D., Botchan M.R., Bouck J., Broksstein P., Brotlier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Domnes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegman C.,
 RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulip D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merklov G., Mlshina N.V., Mobaraj S., Morris J., Mostreli A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wattaman D.A., Weinstein G.M., Weissenbach J., Q.A.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh K.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RA "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003607; AAF52172.1;
 DR Flybase: FBgn0037215; CG12582.
 DR SEQUENCE 347 AA; 39271 MW; F658319B68BA88 CRC64;

Query Match 18.0%; Score 67.5; DB 5; Length 347;
 Best Local Similarity 30.0%; Pred. No. 11;
 Matches 21; Conservative 12; Mismatches 22; Indels 15; Gaps 4;

OY 6 DLSADPAGATAYASKDLEQRLSVSSVD---MTVLYPEYWKY-KCOLRKGSMQH 61
 DB 202 DWGPAPASG-----TKNNLEIIEYAVIREVDVSRVNGSYWNHITGCFDAGQWQ- 255
 OY 62 NREQANLNSR 71
 DB 256 -----NENGR 260

RESULT 8
 O9J2Y9 PRELIMINARY; PRT; 453 AA.
 AC O9J2Y9;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE POLYA POLYMERASE.
 GN NMB0843.
 OS *Neisseria meningitidis* (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / SEROGROUP B;
 RX MEDLINE=2017575; PubMed=10710307;
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson K.J.,
 RA Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Maignani V., Pizzo M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT Complete genome sequence of *Neisseria meningitidis* serogroup B strain

RT MC58.";
 RL Science 287:1809-1815(2000).
 DR EMBL: AE002437; AAF41254.1;
 DR TIGR: NMB0843.
 DR InterPro: IPR001871;
 DR InterPro: IPR002646;
 DR Pfam: PF01743; Polya.pol.1.
 DR PROSITE: PS00036; BZIP_BASIC; UNKNOWN_1.
 DR SEQUENCE 453 AA; 51654 MW; 0BE14F3EB63F7212 CRC64;

Query Match 17.9%; Score 67; DB 2; Length 453;
 Best Local Similarity 27.5%; Pred. No. 17;
 Matches 19; Conservative 16; Mismatches 20; Indels 14; Gaps 2;

OY 3 SGDLSD-----AEPDAGEATAYASKDLEQRLSVSSVD---ELMTVYPEYWK 48
 DB 252 NGFDIPDDHPLMLNRVSDGIAKMTVIALKNTDERLRADKSVSGVFLAALMPELER 311
 OY 49 MYKCOLRKG 57
 DB 312 HMKSNLQOG 320

RESULT 9
 O9JY08 PRELIMINARY; PRT; 453 AA.
 AC O9JY08;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE PUTATIVE POLY(A) POLYMERASE (EC 2.7.7.19).
 GN PCNB OR NMA1053.
 OS *Neisseria meningitidis* (serogroup A).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=65699;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=22491 / SEROTYPE 4A;
 RX MEDLINE=2022556; PubMed=10761919;
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holtroyd S.,
 RA Jagsels K., Leather S., Moule S., Mungall K., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrall B.G.;
 RT "Complete DNA sequence of a serogroup A strain of *Neisseria*
 RT meningitidis 22491.";
 RL Nature 404:502-506(2000).
 DR EMBL: AL162755; CAB84318.1;
 DR InterPro: IPR001871;
 DR InterPro: IPR002646;
 DR Pfam: PF01743; Polya.pol.1.
 DR PROSITE: PS00036; BZIP_BASIC; UNKNOWN_1.
 DR Transference: Nucleotidyltransferase.
 DR SEQUENCE 453 AA; 51719 MW; FCBB4E93F6274DA8 CRC64;

Query Match 17.9%; Score 67; DB 2; Length 453;
 Best Local Similarity 27.5%; Pred. No. 17;
 Matches 19; Conservative 16; Mismatches 20; Indels 14; Gaps 2;

OY 3 SGDLSD-----AEPDAGEATAYASKDLEQRLSVSSVD---ELMTVYPEYWK 48
 DB 252 NGFDIPDDHPLMLNRVSDGIAKMTVIALKNTDERLRADKSVSGVFLAALMPELER 311
 OY 49 MYKCOLRKG 57
 DB 312 HMKSNLQOG 320

RESULT 10
 O9JSTEO

ID Q9STE0 PRELIMINARY; PRT; 1299 AA.
 AC Q9STE0;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE MAP3 EPSILON 1 PROTEIN KINASE.
 GN MAP3KEL.
 OS Brassica napus (Rape).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3708;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. TOPAS. TISSUE=MICROSPORE-DERIVED EMBRYOS;
 RA Jouanne S., Tregear J.W., Kreis M., Henry Y.;
 RT "Characterisation of a B. napus cDNA highly similar to the A. thaliana
 RT AtMAP3Kepsilon1 gene."
 RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AJ238845; CAB54520.1; -
 DR HSP; 000534; 1817.
 DR InterPro: IPR000719; -
 DR InterPro: IPR001245; -
 DR Pfam: PFO0069; pkinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR SMART: SM00220; S_TKC; 1.
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase;
 KW Tyrosine-protein kinase.
 SO SEQUENCE 1299 AA; 143653 MW; 503E946C0A05358F CRC64;

Query Match 17.9%; Score 67; DB 10; Length 1299;
 Best Local Similarity 41.9%; Pred. No. 61;
 Matches 13; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

QY 19 AYASKLEQLSVSSVDLMVLPYKMK 49
 DB 1117 AHASRNSREQLSHGGLDYLSDLEDEYMSV 1147

RESULT 11
 Q9K397 PRELIMINARY; PRT; 602 AA.
 AC Q9K397;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE CG7246 PROTEIN.
 GN CG7246.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Branton R.C., Rogers J.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,

RA Butlis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher S., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,
 RA Jialili M.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003446; AAF46437.1; -
 DR FlyBase: FBgn0030081; CG7246.
 DR InterPro: IPR003107; -
 DR SMART: SM00386; HAT; 1.
 SO SEQUENCE 602 AA; 70103 MW; 2BB8EB9A040FCFEE CRC64;

Query Match 17.6%; Score 66; DB 5; Length 602;
 Best Local Similarity 31.5%; Pred. No. 32;
 Matches 17; Conservative 11; Mismatches 18; Indels 8; Gaps 2;

QY 2 EGGDLSDAEPDAGCETAYASKLEQLRSV-----SSVDLMVLPYKMKY 50
 DB 283 EBDLDTSDDEPASKRSRSYNGRSLKRIOLCVTKYKSAVEELQT---EMNNKY 333

RESULT 12
 Q9K022 PRELIMINARY; PRT; 703 AA.
 AC Q9K022;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DE ATP-DEPENDENT HELICASE, DING FAMILY.
 GN VC1855.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Enomoleva M.D., Vamathevan J., Bess S., Qin H., Dragoti I., Sellers P.,
 RA McDonald L., Ullrich T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae."
 RL Nature 406:477-483(2000).
 DR EMBL: AE004261; AAF95003.1; -
 DR TIGR: VC1855; -
 DR InterPro: IPR001989; -

DR Prodrom: PD004758; -: 1.
KW Helicase.
SQ SEQUENCE 703 AA; 78488 MW; FC88B361B8A8578C CRC64;

Query Match
Best Local Similarity 17.6%; Score 66; DB 2; Length 703;
Matches 18; Conservative 9; Mismatches 20; Indels 10; Gaps 2;

OY 9 DAEPDAGEATAYASKDLEBOLRSVSSVDLMVLYPEYWM-----YKCOLKRGCM 59
Db 514 EMEPQAGEATAYAKKVLCLYQ-----DKANLVLFASYQMREVASLAEVETKRCGM 566

RESULT 13
O9YQ02 PRELIMINARY; PRT; 1612 AA.

AC O9YQ02:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE CG2779 PROTEIN.
GN CG2779.

OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Cealinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Suton G.G., Wortman J.R., Yandell M.D., Zheng Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Duthin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Meltel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Neufuss D.R., Nelson K.A., Nixon R., Nusskern D.R., Pacle J.M.,
RA Palazozo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Schebler F., Shen H.,
RA Shue B.C., Sidenkimos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003488; AAF48140.2; -
DR FlyBase; FBgn0030371; CG2779.
SQ SEQUENCE 1612 AA; 149407 MW; D73877D13EF5F24C CRC64;

Query Match
Best Local Similarity 17.3%; Score 65; DB 5; Length 1612;
Matches 18; Conservative 10; Mismatches 19; Indels 6; Gaps 3;

OY 9 DAEPDAGEATAYASKDLEBOLRSVSSVDLMVLYPEY---WKMYCOLKRGCM 58
Db 1530 DDEPSSKEKKQYV-KDLIRKLKKGDECD--DDVYPDVRCRRYRYCEVKKSG 1579

RESULT 14
O22184 PRELIMINARY; PRT; 2408 AA.

AC O22184:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE T05A1.4 PROTEIN.
GN T05A1.4.

OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]

RP SEQUENCE FROM N.A.
RL Lloyd C.;
RP Submitted (DEC-1995) to the EMBL/Genbank/DBJ databases.

RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton J.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
DR EMBL; 268219; CAA92477.1; -
DR InterPro; IPR001584; -
DR InterPro; IPR001969; -
DR InterPro; IPR002106; -
DR Pfam; PF00665; IVE; 1.
DR PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 2408 AA; 269763 MW; 6C9542E1FA80AC09 CRC64;

Query Match
Best Local Similarity 17.3%; Score 65; DB 5; Length 2408;
Matches 25; Conservative 10; Mismatches 27; Indels 24; Gaps 5;

OY 8 SDAEPDAGEATAYASKDLE-----EQLRVSSVD-----FLMVLY---YREY 46
Db 1530 TDASODITAAAYAAHFKESWAPVTRLIAKNNKIKETISVTYITPILKLLTIRTSRTE 1589

OY 47 WKMYCOLKRGCM-GWQHNH-EQANLN 69
Db 1590 QKSFACRSRKKRAVYHNRVQYNNEN 1615

RESULT 15
O9QZ01 PRELIMINARY; PRT; 4488 AA.

AC O9QZ01:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, last annotation update)
 DE LEFT-RIGHT DYNEIN.
 GN DNAC11 OR LRD.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA SUPP D.M., Brueckner M., Kuehn M.R., Witte D.P., Lowe L.A.,
 RA McGrath J., Corrales J.M., Potter S.S.;
 RT "Targeted deletion of the ATP binding domain of left-right dynein
 RT confirms its role in specifying development of left-right
 RT asymmetries.";
 RL Development 0:0-0(1999).
 DR EMBL: AF183144; AAF07922.1; -;
 DR MGD: MG1:1100864; Dnacc11.
 DR InterPro: IPR00169; -;
 DR InterPro: IPR001680; -;
 DR InterPro: IPR001993; -;
 DR InterPro: IPR003593; -;
 DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
 DR PROSITE: PS00639; THIOL_PROTEASE_HTS; UNKNOWN_1.
 DR PROSITE: PS00678; WD_REPEATS; UNKNOWN_1.
 DR SMART: SM00382; AAA; 1.
 SQ SEQUENCE 4488 AA; 516184 MW; 1C5E0050928D949A CRC64;

Query Match 17.28; Score 64.5; DB 11; Length 4488;
 Best Local Similarity 26.7%; Pred. No. 5.1e+02;
 Matches 24; Conservative 11; Mismatches 22; Indels 33; Gaps 3;

OY 5 LDISDAEPDAGEATAVASRKLEEQLSRSVSDLEMTLYPEYKMYKC-----QLRKGW- 59
 DB 1263 LDKANQLELA-----LEEBMEQONSARKFEVALPEYKMYKC-----QLRKGW- 59
 OY 60 -----QHNRQANLNSR 71
 DB 1313 VIIVRRSIDNMTETQWQINVEQMDLELR 1342

Search completed: October 17, 2001, 14:50:55
 Job time: 421 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:45:59 ; Search time 115.93 Seconds
(without alignments)
109.294 Million cell updates/sec

Title: US-09-427-657-2_COPY_211_419
Perfect score: 1218
Sequence: 1 CSMKLDYRQVHSHIRSLP.....SYSEVCRCPVYMKRPQMS 209

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.GeneSeq_0601.*
1: /SID88/gcgdata/geneSeq/geneSeq/AA1980.DAT.*
2: /SID88/gcgdata/geneSeq/geneSeq/AA1981.DAT.*
3: /SID88/gcgdata/geneSeq/geneSeq/AA1982.DAT.*
4: /SID88/gcgdata/geneSeq/geneSeq/AA1983.DAT.*
5: /SID88/gcgdata/geneSeq/geneSeq/AA1984.DAT.*
6: /SID88/gcgdata/geneSeq/geneSeq/AA1985.DAT.*
7: /SID88/gcgdata/geneSeq/geneSeq/AA1986.DAT.*
8: /SID88/gcgdata/geneSeq/geneSeq/AA1987.DAT.*
9: /SID88/gcgdata/geneSeq/geneSeq/AA1988.DAT.*
10: /SID88/gcgdata/geneSeq/geneSeq/AA1989.DAT.*
11: /SID88/gcgdata/geneSeq/geneSeq/AA1990.DAT.*
12: /SID88/gcgdata/geneSeq/geneSeq/AA1991.DAT.*
13: /SID88/gcgdata/geneSeq/geneSeq/AA1992.DAT.*
14: /SID88/gcgdata/geneSeq/geneSeq/AA1993.DAT.*
15: /SID88/gcgdata/geneSeq/geneSeq/AA1994.DAT.*
16: /SID88/gcgdata/geneSeq/geneSeq/AA1995.DAT.*
17: /SID88/gcgdata/geneSeq/geneSeq/AA1996.DAT.*
18: /SID88/gcgdata/geneSeq/geneSeq/AA1997.DAT.*
19: /SID88/gcgdata/geneSeq/geneSeq/AA1998.DAT.*
20: /SID88/gcgdata/geneSeq/geneSeq/AA1999.DAT.*
21: /SID88/gcgdata/geneSeq/geneSeq/AA2000.DAT.*
22: /SID88/gcgdata/geneSeq/geneSeq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	1218	100.0	292 20 AAW86225	Human VEGF-C trunc
2	1218	100.0	297 20 AAW86224	Human VEGF-C trunc
3	1218	100.0	302 20 AAW86223	Human VEGF-C trunc
4	1218	100.0	307 20 AAW86222	Human VEGF-C trunc
5	1218	100.0	318 20 AAW86221	Human VEGF-C trunc
6	1218	100.0	399 20 AAW86237	Human VEGF-C full
7	1218	100.0	419 18 AAW17837	Human foetal liver
8	1218	100.0	419 18 AAW0932	Human Flt4 recepto
9	1218	100.0	419 18 AAW13833	Human vascular end
10	1218	100.0	419 19 AAW5751	Vascular endotheli
11	1218	100.0	419 19 AAW57540	Human vascular end

12	1218	100.0	419 20 AAW86203	Human vascular end
13	1218	100.0	419 21 AAB10648	Human VEGF-C prote
14	1218	100.0	419 21 AAB28048	Human VEGF-C prote
15	1218	100.0	419 21 AAV70749	Human prepro-vascu
16	1218	100.0	419 21 AAV70982	Human vascular end
17	1218	100.0	419 22 AAB37605	Human VEGF-C. Hom
18	1214	99.7	350 16 AAB82686	Vascular endotheli
19	1214	99.7	350 20 AAV30519	A truncated vascul
20	1214	99.7	350 20 AAY23221	Truncated human VE
21	1214	99.7	350 21 AAV97145	Truncated vascular
22	1214	99.7	350 22 AAV97577	Human VEGF-2 prote
23	1214	99.7	419 20 AAV30518	Vascular endotheli
24	1214	99.7	419 20 AAV22320	Full length human
25	1214	99.7	419 21 AAV97144	Vascular endotheli
26	1214	99.7	419 22 AAV97570	Human VEGF-B prote
27	1211	99.4	419 18 AAW11478	Human vascular end
28	1086	89.2	415 18 AAW00933	Mouse Flt4 recepto
29	1086	89.2	415 19 AAW57442	Mouse vascular end
30	927	76.1	418 18 AAW00934	Ouail Flt4 recepto
31	927	76.1	418 19 AAW5743	Ouail vascular end
32	253.5	20.8	325 19 AAW53240	Homo sapiens vascu
33	253.5	20.8	325 22 AAV97572	Human VEGF-D prote
34	253.5	20.8	354 19 AAW49036	Human zveig12 growt
35	253.5	20.8	354 19 AAW53241	Homo sapiens vascu
36	253.5	20.8	354 19 AAW44293	Human vascular end
37	253.5	20.8	354 21 AAB10649	Human VEGD protein
38	253.5	20.8	354 21 AAB28049	Human VEGF-D prote
39	253.5	20.8	354 21 AAV70750	Human prepro-vascu
40	253.5	20.8	354 21 AAV70983	Human vascular end
41	253.5	20.8	354 22 AAV97573	Human VEGF-D1 prot
42	253.5	20.8	354 22 AAB37606	Human VEGF-D. Hom
43	249.5	20.5	354 22 AAB70685	Human vascular end
44	241.5	19.8	337 20 AAV08286	Human growth facto
45	241.5	19.8	338 18 AAW14992	Marline c-Fos induc

ALIGNMENTS

RESULT 1
AAW86225 standard; protein; 292 aa.
ID AAW86225;
AC AAW86225;
XX 16-FEB-1999 (first entry)
XX
DT Human VEGF-C truncated fragment 4.
DE
XX VEGF: VEGF: vascular endothelial growth factor; VEGF-related protein;
XX recombinant; truncated; gene therapy; angiogenesis; cardiac ischaemia;
XX coronary; collateral vessel development; cell growth; migration; heart;
XX lower limb ischaemia; stroke; peripheral vascular disease; intestine;
XX wound healing; skin; vascular permeability.
OS Homo sapiens.
XX
XX W09849300-A2.
XX
XX 05-NOV-1998.
XX
XX 20-APR-1998; 98WO-US07801.
XX
XX 25-APR-1997; 97US-0842984.
XX
XX (COLL-) COLLATERAL THERAPEUTICS.
XX
XX Bohlen P;
XX WPI: 1999-009426/01.
XX
XX New truncated vascular endothelial growth factor-related protein
XX subunits - lack part of the N-terminal sequence, used to stimulate

PT angiogenesis, e.g. for treating heart disease and ischaemia
XX
PS Claim 5; Fig 2D; 113pp; English.
XX
CC The invention relates to truncated VRP (vascular endothelial growth
CC factor (VEGF)-related protein) subunits that have at least one amino
CC acid N-terminal to the first Cys of the core sequence deleted. Host
CC cells transfected or transfected with expression vectors containing
CC nucleic acids encoding the truncated VRP subunits are used to produce
CC the truncated proteins recombinantly. The truncated VRP subunits,
CC optionally expressed from gene therapy vectors, have in vivo and in vitro
CC angiogenic activity and are used to stimulate angiogenesis, particularly
CC coronary collateral vessel development in cases of cardiac ischaemia; to
CC stimulate endothelial cell growth and migration in vitro; to treat heart
CC disease; to treat ischaemia (e.g. cardiac, chronic coronary or chronic
CC lower limb ischaemia; stroke and peripheral vascular disease); to promote
CC healing of wounds (of skin or intestines), and to increase vascular
CC permeability. Sequences AAW86222 to AAW86225 represent truncated
CC fragments of VEGF-C.
XX
SQ Sequence 292 AA:

Query Match 100.0%; Score 1218; DB 20; Length 292;
Best Local Similarity 100.0%; Pred. No. 4.3e-90;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CMSKLDVYRQVSHIIRSRIPATLPQCAANKTCPTNYMNMNHICRLAEDFMFSSDAGD 60
DB 84 cmskldvyrvqshislrspatlpqcaanktcptnymnmnhicrlagedfmfssdagd 143
OY 61 DSTDGFHDICGNKELDEFTCCQVCNAGLRPASCGPHKELDRNSCQVCYKKNLFPSQCGA 120
DB 144 dstdgfhndicgnkeldetccqvcnaglrpsacgphkeldrnsqcvcyknkllfpsqga 203
OY 121 NRPEDNTQCVCYKRCRCPNPNPLNGKCKACECTESPQKLLGKKFHHQTCGYRRPCIN 180
DB 204 nreidentcvcykcrtcpnplngkckacectespqkllgkfhhtqtcgyrrpcin 263
OY 181 ROKACEPGFSYEVCRCVPYSYWKRRPOMS 209
DB 264 rkaacepgfisyeevrcrcvpsywkrrpoms 292

RESULT 2
AAW86224
ID AAW86224 standard; protein; 297 AA.
XX
AC AAW86224;
XX
DT 16-FEB-1999 (first entry)
XX
DE Human VEGF-C truncated fragment 3.
XX
XX VEGF: VRP; vascular endothelial growth factor; VEGF-related protein;
KM recombinant; truncated; gene therapy; angiogenesis; cardiac ischaemia;
KM coronary; collateral vessel development; cell growth; migration; heart;
KM lower limb ischaemia; stroke; peripheral vascular disease; intestine;
KM wound healing; skin; vascular permeability.
XX
OS Homo sapiens.
XX
PN WO9849300-A2.
XX
PD 05-NOV-1998.
XX
PE 20-APR-1998; 98WO-US07801.
XX
PR 25-APR-1997; 97US-0842984.
XX
PA (COLL-) COLLATERAL THERAPEUTICS.
XX
PI Bohlen P;

XX
DR WPI; 1999-009426/O1.
XX
PT New truncated vascular endothelial growth factor-related protein
PT subunits - lack part of the N-terminal sequence, used to stimulate
PT angiogenesis, e.g. for treating heart disease and ischaemia
XX
PS Claim 5; Fig 2D; 113pp; English.
XX
CC The invention relates to truncated VRP (vascular endothelial growth
CC factor (VEGF)-related protein) subunits that have at least one amino
CC acid N-terminal to the first Cys of the core sequence deleted. Host
CC cells transfected or transfected with expression vectors containing
CC nucleic acids encoding the truncated VRP subunits are used to produce
CC the truncated proteins recombinantly. The truncated VRP subunits,
CC optionally expressed from gene therapy vectors, have in vivo and in vitro
CC angiogenic activity and are used to stimulate angiogenesis, particularly
CC coronary collateral vessel development in cases of cardiac ischaemia; to
CC stimulate endothelial cell growth and migration in vitro; to treat heart
CC disease; to treat ischaemia (e.g. cardiac, chronic coronary or chronic
CC lower limb ischaemia; stroke and peripheral vascular disease); to promote
CC healing of wounds (of skin or intestines), and to increase vascular
CC permeability. Sequences AAW86222 to AAW86225 represent truncated
CC fragments of VEGF-C.
XX
SQ Sequence 297 AA:

Query Match 100.0%; Score 1218; DB 20; Length 297;
Best Local Similarity 100.0%; Pred. No. 4.3e-90;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CMSKLDVYRQVSHIIRSRIPATLPQCAANKTCPTNYMNMNHICRLAEDFMFSSDAGD 60
DB 89 cmskldvyrvqshislrspatlpqcaanktcptnymnmnhicrlagedfmfssdagd 148
OY 61 DSTDGFHDICGNKELDEFTCCQVCNAGLRPASCGPHKELDRNSCQVCYKKNLFPSQCGA 120
DB 149 dstdgfhndicgnkeldetccqvcnaglrpsacgphkeldrnsqcvcyknkllfpsqga 208
OY 121 NRPEDNTQCVCYKRCRCPNPNPLNGKCKACECTESPQKLLGKKFHHQTCGYRRPCIN 180
DB 209 nreidentcvcykcrtcpnplngkckacectespqkllgkfhhtqtcgyrrpcin 268
OY 181 ROKACEPGFSYEVCRCVPYSYWKRRPOMS 209
DB 269 rkaacepgfisyeevrcrcvpsywkrrpoms 297

RESULT 3
AAW86223
ID AAW86223 standard; protein; 302 AA.
XX
AC AAW86223;
XX
DT 16-FEB-1999 (first entry)
XX
DE Human VEGF-C truncated fragment 2.
XX
XX VEGF: VRP; vascular endothelial growth factor; VEGF-related protein;
KM recombinant; truncated; gene therapy; angiogenesis; cardiac ischaemia;
KM coronary; collateral vessel development; cell growth; migration; heart;
KM lower limb ischaemia; stroke; peripheral vascular disease; intestine;
KM wound healing; skin; vascular permeability.
XX
OS Homo sapiens.
XX
PN WO9849300-A2.
XX
PD 05-NOV-1998.
XX
PE 20-APR-1998; 98WO-US07801.
XX

PR 25-APR-1997; 97US-0842984.
XX (COLL-) COLLATERAL THERAPEUTICS.
XX Bohlen P;
XX WPI; 1999-009426/01.
XX New truncated vascular endothelial growth factor-related protein
PT subunits - lack part of the N-terminal sequence, used to stimulate
PT angiogenesis, e.g. for treating heart disease and ischaemia
XX
XX
PS Claim 5; Fig 2D; 113pp; English.
XX The invention relates to truncated VRP (vascular endothelial growth
CC factor (VEGF)-related protein) subunits that have at least one amino
CC acid N-terminal to the first Cys of the core sequence deleted. Host
CC cells transformed or transfected with expression vectors containing
CC nucleic acids encoding the truncated VRP subunits are used to produce
CC the truncated proteins recombinantly. The truncated VRP subunits,
CC optionally expressed from gene therapy vectors, have in vivo and in vitro
CC angiogenic activity and are used to stimulate angiogenesis, particularly
CC coronary collateral vessel development in cases of cardiac ischaemia; to
CC stimulate endothelial cell growth and migration in vitro; to treat heart
CC disease; to treat ischaemia (e.g. cardiac, chronic coronary or chronic
CC lower limb ischaemia; stroke and peripheral vascular disease); to promote
CC healing of wounds (of skin or intestines), and to increase vascular
CC permeability. Sequences AAM86222 to AAM86225 represent truncated
CC fragments of VEGF-C.
XX
XX Sequence 302 AA;
S0
Query Match 100.0%; Score 1218; DB 20; Length 302;
Best Local Similarity 100.0%; Pred. No. 4.4e-90;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CSMKIDVRYOVHSITRSLPATLPQCOAANKTCPTNMMNNHICRCLAOEDPFSSDGD 60
DB 94 cmskldvryovhsitrrslpatlpqcgaaanktcptnmmnhicclaqedfmsdgd 153
OY 61 DSTDFHIDICGNKELDEETQCVCVACGLRPASCGPHKELDRNSQCVCKNKLFPSSQGA 120
DB 154 dstdghidicgnkeldetecgcveraglrpascgphkeldrnsqcvcknklfsgsga 213
OY 121 NREPDENTCQCVCKRTCRNQPPLNPGKACACECTESPQCKLKGKFFHQTSCVRRPCTN 180
DB 214 nreidentcgcvcckrtcrnpplnpgkacacectespqckllgkffhqtscvrrpctn 273
OY 181 ROKACEPGFSYSEVCRCVPSYWKRPOMS 209
DB 274 rkacepqfsysevcrcvpsywkrpoms 302
RESULT 4
AAM86222
ID AAM86222 standard; protein: 307 AA.
XX
XX AAM86222;
XX 16-FEB-1999 (first entry)
XX Human VEGF-C truncated fragment 1.
XX
XX VEGF; VRP; vascular endothelial growth factor; VEGF-related protein;
KW recombinant; truncated; gene therapy; angiogenesis; cardiac ischaemia;
KW coronary; collateral vessel development; cell growth; migration; heart;
KW lower limb ischaemia; stroke; peripheral vascular disease; intestine;
KW wound healing; skin; vascular permeability.
XX
XX Homo sapiens.
OS
XX
PN W09849300-A2.

XX 05-NOV-1998.
PD
XX 20-APR-1998; 98WO-US07801.
PE
XX 25-APR-1997; 97US-0842984.
PR
XX (COLL-) COLLATERAL THERAPEUTICS.
XX Bohlen P;
XX WPI; 1999-009426/01.
XX
XX
PS Claim 5; Fig 2D; 113pp; English.
XX The invention relates to truncated VRP (vascular endothelial growth
CC factor (VEGF)-related protein) subunits that have at least one amino
CC acid N-terminal to the first Cys of the core sequence deleted. Host
CC cells transformed or transfected with expression vectors containing
CC nucleic acids encoding the truncated VRP subunits are used to produce
CC the truncated proteins recombinantly. The truncated VRP subunits,
CC optionally expressed from gene therapy vectors, have in vivo and in vitro
CC angiogenic activity and are used to stimulate angiogenesis, particularly
CC coronary collateral vessel development in cases of cardiac ischaemia; to
CC stimulate endothelial cell growth and migration in vitro; to treat heart
CC disease; to treat ischaemia (e.g. cardiac, chronic coronary or chronic
CC lower limb ischaemia; stroke and peripheral vascular disease); to promote
CC healing of wounds (of skin or intestines), and to increase vascular
CC permeability. Sequences AAM86222 to AAM86225 represent truncated
CC fragments of VEGF-C.
XX
XX Sequence 307 AA;
S0
Query Match 100.0%; Score 1218; DB 20; Length 307;
Best Local Similarity 100.0%; Pred. No. 4.5e-90;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CSMKIDVRYOVHSITRSLPATLPQCOAANKTCPTNMMNNHICRCLAOEDPFSSDGD 60
DB 99 cmskldvryovhsitrrslpatlpqcgaaanktcptnmmnhicclaqedfmsdgd 158
OY 61 DSTDFHIDICGNKELDEETQCVCVACGLRPASCGPHKELDRNSQCVCKNKLFPSSQGA 120
DB 159 dstdghidicgnkeldetecgcveraglrpascgphkeldrnsqcvcknklfsgsga 218
OY 121 NREPDENTCQCVCKRTCRNQPPLNPGKACACECTESPQCKLKGKFFHQTSCVRRPCTN 180
DB 219 nreidentcgcvcckrtcrnpplnpgkacacectespqckllgkffhqtscvrrpctn 278
OY 181 ROKACEPGFSYSEVCRCVPSYWKRPOMS 209
DB 279 rkacepqfsysevcrcvpsywkrpoms 307
RESULT 5
AA08284
ID AA08284 standard; protein: 318 AA.
XX
XX AA08284;
XX 14-JUL-1999 (first entry)
XX Human growth factor protein fragment VEGF-C.
XX
XX Growth factor; human; dimer; cysteine knot; cellular inclusion body;
KW pharmaceutical.
KW
XX
OS
XX
XX Homo sapiens.

XX DE19748734-A1.
 XX
 XX PD 06-MAY-1999.
 XX
 XX PF 05-NOV-1997; 97DE-1048734.
 XX
 XX PR 05-NOV-1997; 97DE-1048734.
 XX
 XX PA (GBFB) GES BIOTECHNOLOGISCHE FORSCHUNG MBH.
 XX
 XX PI Erdmann H, Kaerst U, Mueller C, Rinas U, Weich H;
 XX
 XX DR WPI; 1999-278785/24.
 XX
 XX PT Preparing active growth factor dimers from inclusion bodies in high
 XX yield
 XX
 XX PS Claim 14; Page 8; 14pp; German.
 XX
 XX CC This invention describes the novel preparation of biologically active
 XX dimers of recombinant human growth factors of the cysteine knot family
 XX starting from cellular inclusion bodies. Such dimers are are useful in
 XX pharmaceutical compositions and the method provides yields of 31-39.7%,
 XX in examples, compared with about 10% for the conventional method (see
 XX CC Biochemistry, 28 (1989) 2956). AAY08278-Y08301 are human growth factor
 XX protein fragments used in the method of the invention.
 XX
 XX SQ Sequence 318 AA;

Query Match 100.0%; Score 1218; DB 20; Length 318;
 Best Local Similarity 100.0%; Pred. No. 4.6e-90;
 Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CMSKIDYRVQVHSITRSLPATLPQCOAANKTCPTNYMNNHICRLAOEDMFSSDAGD 60
 DB 110 CMSKIDYRVQVHSITRSLPATLPQCOAANKTCPTNYMNNHICRLAOEDMFSSDAGD 169
 QY 61 DSTDFHHDICGNKELEDETCQVCRCRAGLRPASCPHKELDNSQCVCYKKNLFPSCGA 120
 DB 170 DSTDFHHDICGNKELEDETCQVCRCRAGLRPASCPHKELDNSQCVCYKKNLFPSCGA 229
 QY 121 NREPDENTQCCVCKRTCPRNQPLNPGKACACECTESPQCKLKGKFFHHQTCSCYRRPCTN 180
 DB 230 NREPDENTQCCVCKRTCPRNQPLNPGKACACECTESPQCKLKGKFFHHQTCSCYRRPCTN 289
 QY 181 RQACEPGFSYSEEVCRCPVSYWKRPM 209
 DB 290 RQACEPGFSYSEEVCRCPVSYWKRPM 318

RESULT 6
 AAW86237
 ID AAW86237 standard; protein; 399 AA.
 XX
 XX AC AAW86237;
 XX
 XX DT 16-FEB-1999 (first entry)
 XX
 XX DE Human VEGF-C full length sequence.
 XX
 XX VEGF; VEGF; vascular endothelial growth factor; VEGF-related protein;
 KW recombinant; truncated; gene therapy; angiogenesis; cardiac ischaemia;
 KW coronary; collateral vessel development; cell growth; migration; heart;
 KW lower limb ischaemia; stroke; peripheral vascular disease; intestine;
 KW wound healing; skin; vascular permeability.
 XX
 XX OS Homo sapiens.
 XX
 XX PN WO9849300-A2.
 XX
 XX PD 05-NOV-1998.

XX 20-APR-1998; 98WO-US07801.
 XX
 XX XX 25-APR-1997; 97US-0842984.
 XX
 XX PA (COLL-) COLLATERAL THERAPEUTICS.
 XX
 XX PI Bohlen P;
 XX
 XX DR WPI; 1999-009426/01.
 XX
 XX PT New truncated vascular endothelial growth factor-related protein
 XX subunits - lack part of the N-terminal sequence, used to stimulate
 XX PT angiogenesis, e.g. for treating heart disease and ischaemia
 XX
 XX PS Claim 5; Fig 2D; 113pp; English.

The invention relates to truncated VRP (vascular endothelial growth
 factor (VEGF)-related protein) subunits that have at least one amino
 acid N-terminal to the first Cys of the core sequence deleted. Host
 cells transformed or transfected with expression vectors containing
 nucleic acids encoding the truncated VRP subunits are used to produce
 the truncated proteins recombinantly. The truncated VRP subunits,
 optionally expressed from gene therapy vectors, have in vivo and in vitro
 angiogenic activity and are used to stimulate angiogenesis, particularly
 coronary collateral vessel development in cases of cardiac ischaemia; to
 stimulate endothelial cell growth and migration in vitro; to treat heart
 disease; to treat ischaemia (e.g. cardiac, chronic coronary or chronic
 lower limb ischaemia; stroke and peripheral vascular disease); to promote
 healing of wounds (of skin or intestines), and to increase vascular
 permeability. Sequences AAW86234 to AAW86239 represent full length VRP
 sequences from which the truncated fragments are created.

Query Match 100.0%; Score 1218; DB 20; Length 399;
 Best Local Similarity 100.0%; Pred. No. 5.7e-90;
 Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CMSKIDYRVQVHSITRSLPATLPQCOAANKTCPTNYMNNHICRLAOEDMFSSDAGD 60
 DB 191 CMSKIDYRVQVHSITRSLPATLPQCOAANKTCPTNYMNNHICRLAOEDMFSSDAGD 250
 QY 61 DSTDFHHDICGNKELEDETCQVCRCRAGLRPASCPHKELDNSQCVCYKKNLFPSCGA 120
 DB 251 DSTDFHHDICGNKELEDETCQVCRCRAGLRPASCPHKELDNSQCVCYKKNLFPSCGA 310
 QY 121 NREPDENTQCCVCKRTCPRNQPLNPGKACACECTESPQCKLKGKFFHHQTCSCYRRPCTN 180
 DB 311 NREPDENTQCCVCKRTCPRNQPLNPGKACACECTESPQCKLKGKFFHHQTCSCYRRPCTN 370
 QY 181 RQACEPGFSYSEEVCRCPVSYWKRPM 209
 DB 371 RQACEPGFSYSEEVCRCPVSYWKRPM 399

RESULT 7
 AAW17837
 ID AAW17837 standard; protein; 419 AA.
 XX
 XX AC AAW17837;
 XX
 XX DT 13-JAN-1998 (first entry)
 XX
 XX DE Human foetal liver kinase A binding protein flk-1bp.
 XX
 XX Foetal liver kinase I binding protein; human; flk-1bp;
 KW receptor tyrosine kinase; vasculogenesis; angiogenesis;
 KW wound healing; tumour; therapy; antagonist; antibody.
 XX
 XX OS Homo sapiens.
 XX
 XX PD

Key	Location/Qualifiers
FH	1..20
FT	/label= Sig-peptide
FT	21..419
FT	/label= Mat-protein
FT	/note= "(Claim 10)"
FT	21..35
FT	/label= N-terminal
FT	/note= "(Claim 9)"
PN	
PD	WO9717442-A1.
XX	15-MAY-1997.
XX	
PF	05-NOV-1996; 96WO-0517584.
XX	
PR	08-NOV-1995; 95US-0554374.
XX	
PA	(IMMV) IMMUNEX CORP.
XX	
P1	Lyman SD;
XX	
DR	WPI: 1997-281031/25.
XX	
DR	N-PSDB: AAT68811.
PT	DNA encoding a human foetal liver kinase 1 binding protein - used
PT	to treat conditions with insufficient protein, deliver agents to
PT	cells and identify antagonists to treat protein-mediated conditions
XX	
PS	Claim 1; Page 30-32; 43pp; English.
XX	
CC	This polypeptide comprises a human foetal liver kinase 1 binding
CC	protein (flk-1bp) (see AAW17837) that binds to the receptor tyrosine
CC	kinase flk-1 expressed on vascular endothelial and other cells.
CC	The mature flk1-bp can be secreted from host cells transformed with
CC	an expression vector including an isolated flk-1bp cDNA clone (see
CC	AAT68811). flk-1bp can be used to isolate cells to which it binds,
CC	for use in studying the roles of such cells and of flk-1 in
CC	vasculogenesis and angiogenesis. Angiogenesis inhibition or
CC	increased vasculature growth or in wound healing, respectively).
CC	The flk-1bp can be administered to treat conditions with defective
CC	or insufficient flk-1. Polypeptides may also act as carriers to
CC	deliver diagnostic/therapeutic agents to cells to which flk1-bp
CC	binds, to generate antibodies, and to identify flk-1bp antagonists
CC	useful for treating flk-1bp mediated conditions.
XX	
XX	Sequence 419 AA:
Query Match	100.0%; Score 1218; DB 18; Length 419;
Best Local Similarity	100.0%; Pred. No. 6e-90;
Matches 209; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
1 CMSKLDYRQVHSTIRRSIPATLPQCOAANKTCPTNYMNNHHCICLAEDTFMSSDAGD	60
211 CMSKLDYRQVHSTIRRSIPATLPQCOAANKTCPTNYMNNHHCICLAEDTFMSSDAGD	270
61 DSTDFEHIDCGNKELDETCOCVCRAGLRPRSCGPRKELDINSOCVCKNKLFPSSGCA	120
271 DSTDFEHIDCGNKELDETCOCVCRAGLRPRSCGPRKELDINSOCVCKNKLFPSSGCA	330
121 NREPDENTCOCVCKRTPCRNOPLINPGKACAEETESPQKCLLGGKFFHHQTCSCYRPRCTN	180
331 NREPDENTCOCVCKRTPCRNOPLINPGKACAEETESPQKCLLGGKFFHHQTCSCYRPRCTN	390
181 ROKACEPGFSTSEEVCRVPSYWKRPQMS	209
391 ROKACEPGFSTSEEVCRVPSYWKRPQMS	419

XX	AAW00932 standard; Protein: 419 AA.
XX	AAW00932;
XX	10-NOV-1997 (first entry)
XX	Human Flt4 receptor tyrosine kinase ligand VEGF-C.
XX	VEGF-C; Flt4; receptor tyrosine kinase; VEGFR-3; human;
XX	vascular endothelial growth factor receptor-3; ligand;
XX	angiogenesis; wound healing; lymph vessel; lymphangio-
XX	cancer; metastasis; therapy; diagnosis; antibody; inhibitor.
XX	Homo sapiens.
XX	Key
XX	Location/Qualifiers
XX	1..102
XX	/label= prepro_peptide
XX	32..227
XX	/note= "preferred active fragment of VEGF-C,"
XX	retaining Flt4 ligand activity (Claim 15)"
XX	103..217
XX	/note= "preferred active fragment of VEGF-C,"
XX	retaining Flt4 ligand activity (Claim 12)"
XX	103..225
XX	/note= "preferred active fragment of VEGF-C,"
XX	retaining Flt4 ligand activity (Claim 13)"
XX	103..227
XX	/note= "preferred active fragment of VEGF-C,"
XX	retaining Flt4 ligand activity (Claim 14)"
XX	113..213
XX	/note= "preferred active fragment of VEGF-C,"
XX	retaining Flt4 ligand activity (Claim 10)"
XX	113..227
XX	/note= "preferred active fragment of VEGF-C,"
XX	retaining Flt4 ligand activity (Claim 11)"
XX	131..211
XX	/note= "preferred active fragment of VEGF-C,"
XX	retaining Flt4 ligand activity (Claim 9)"
XX	161..221
XX	/note= "preferred active fragment of VEGF-C,"
XX	retaining Flt4 ligand activity (Claim 8)"
XX	WC9705250-A2.
XX	13-FEB-1997.
XX	01-AUG-1996; 96WO-FI00427.
XX	28-JUN-1996; 96US-0671573.
XX	01-AUG-1995; 95US-0510133.
XX	12-JAN-1996; 96US-0585895.
XX	14-FEB-1996; 96US-0601132.
XX	(UYHE-) UNIV HELSINKI LICENSING LTD OY.
XX	Alitalo K, Joukov V;
XX	WPI: 1997-145688/13.
XX	N-PSDB: AAT84276.
XX	Flt4 receptor tyrosine kinase ligand and related nucleic acid - used
XX	to modulate growth of endothelial cells and for diagnosis of
XX	endothelial cell diseases
XX	Claim 7; Page 112-113; 183pp; English.
XX	This polypeptide comprises the pre-pro sequence of human VEGF-C,
XX	a novel ligand that binds specifically to human Flt4 receptor
XX	tyrosine kinase (VEGFR-3), stimulating phosphorylation of the
XX	receptor. Its sequence was deduced from a cDNA clone (AAT84276)
XX	cod. from a PC-3 prostatic adenocarcinoma cell (ATCC CRL 1453)
XX	library. The polypeptide, or its active fragments, can be

CC expressed in transformed or transfected host cells for use in
 CC claimed methods for detecting endothelial cells (e.g. to image
 CC lymphatic vessels, endothelial venules, Flt4 receptor in
 CC histochemical tissue) and also to modulate the growth of mammalian
 CC endothelial cells (e.g. to accelerate angiogenesis and to promote
 CC endothelial function of lymphatic vessels). Inhibitors of
 CC VEGF-C, such as antibodies, can be used to control endothelial
 CC cell proliferation, e.g. lymphangioma or metastatic cancer.
 CC Mouse and quail VEGF-C sequences (see AAM0934-35) have also been
 CC isolated.

XX Sequence 419 AA:

Query Match 100.0%; Score 1218; DB 18; Length 419;
 Best Local Similarity 100.0%; Pred. No. 6e-90;
 Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CMSKLDVYRQVHSIIRSLPATLPQCQAANKTCPTNYMNMNHICRLAODEFMFSSDAGD 60
 DB 211 cmskldvyrvhslirslpatlpqcganktcptnymnmnhicrlagedfmfssdagd 270
 QY 61 DSTDGFHDICGPNKELDEETCCQVCRAGLRPASCGPHKELDNSQCCKNKLFPSCGA 120
 DB 271 dstdgfhdcgpnkeldetccqvcraglrpascgphkeldnscqcknkllfpscga 330
 QY 121 NREFDENTCQCVCKRTCPNPOLNPGKCAECTESPQKCLKGKFFHQTCSQYRPTN 180
 DB 331 nrefdentcqcvcrtcpnpglnpgkcaectespqckllkgkffhqtcsqyrptcn 390
 QY 181 RQACEPGFSYSEVCRCPVSYWKRQMS 209
 DB 391 rqacepgfsysevcrcvpvswkrqms 419

RESULT 9

AAM13833
 ID AAM13833 standard; Protein; 419 AA.

AC AAM13833;
 XX
 DT 05-JUN-1997 (first entry)

DE Human vascular endothelial growth factor-related protein VRF.
 XX
 KW Vascular endothelial growth factor-related protein; VRF; VEGF;
 KW receptor protein tyrosine kinase; Flt4; signal transduction;
 KW wound healing; vulnerrary; rheumatoid arthritis; Kaposi's sarcoma;
 KW therapy; diagnosis; angiogenesis; monoclonal antibody.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH Peptide 1..20
 FT /label= Sig_peptide
 FT Protein 20..419
 FT /label= Mat_protein
 FT Misc-difference 114
 FT /note= "deduced residue from nucleotide sequence
 is tyrosine"

XX WO9709427-A1.

XX 13-MAR-1997.

XX 30-AUG-1996; 96WO-US14075.

XX 08-SEP-1995; 95US-0003491.

XX (GETH) GENENTECH INC.

XX Lee J, Wood W;

XX

DR WPI, 1997-192902/17.

DR N-PSDB: AAT59929.

PT Human protein similar to vascular endothelial growth factor - used
 PT to treat e.g. wounds, tumours, rheumatoid arthritis, Kaposi's
 PT sarcoma etc.

PS Claim 6; Fig 1A-D; 68pp; English.

CC A human vascular endothelial growth factor (VEGF)-related protein
 CC (VRF) (AAM13833) has been identified that binds to, and stimulates
 CC the phosphorylation of, the receptor tyrosine kinase Flt4. It is
 CC postulated to be a third member of the VEGF protein family. Its
 CC amino acid sequence was deduced from a cDNA clone (AAT59929) obtd.
 CC from a glioma G61 library. Recombinant VRF can be produced in
 CC transformed host cells and used: to promote growth of vascular and
 CC lymph endothelial cells; to stimulate phosphorylation of the
 CC tyrosine kinase domain of a Flt4 receptor; as a diagnostic; as an
 CC additive to cell cultures; to screen for (ant)agonists; and to
 CC raise monoclonal antibodies used to treat conditions associated
 CC with excessive neovascularisation or vascular permeability. VRF
 CC may make it possible to avoid coronary by-pass surgery by
 CC stimulating growth of the collateral circulation.

XX Sequence 419 AA:

Query Match 100.0%; Score 1218; DB 18; Length 419;
 Best Local Similarity 100.0%; Pred. No. 6e-90;
 Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CMSKLDVYRQVHSIIRSLPATLPQCQAANKTCPTNYMNMNHICRLAODEFMFSSDAGD 60
 DB 211 cmskldvyrvhslirslpatlpqcganktcptnymnmnhicrlagedfmfssdagd 270
 QY 61 DSTDGFHDICGPNKELDEETCCQVCRAGLRPASCGPHKELDNSQCCKNKLFPSCGA 120
 DB 271 dstdgfhdcgpnkeldetccqvcraglrpascgphkeldnscqcknkllfpscga 330
 QY 121 NREFDENTCQCVCKRTCPNPOLNPGKCAECTESPQKCLKGKFFHQTCSQYRPTN 180
 DB 331 nrefdentcqcvcrtcpnpglnpgkcaectespqckllkgkffhqtcsqyrptcn 390
 QY 181 RQACEPGFSYSEVCRCPVSYWKRQMS 209
 DB 391 rqacepgfsysevcrcvpvswkrqms 419

RESULT 10

AAM75751
 ID AAM75751 standard; Protein; 419 AA.

AC AAM75751;

XX 14-DEC-1998 (first entry)

DE Vascular endothelial growth factor C protein analogue.

KW Flt4; vascular endothelial growth factor C; vascular endothelial cell;
 KW lymphatic endothelial cell; myelopoiesis; angiogenesis; inflammation;
 KW lymphangiogenesis; oedema; elephantiasis; Milroy's disease.

XX Synthetic.

XX Key Location/Qualifiers
 FH Modified-site 156
 FT /note= "xaa can be anything other than cysteine, or
 FT can be nothing"

XX WO9833917-A1.

XX 06-AUG-1998.

XX

PF 02-FEB-1998; 98WO-US01973.
 XX
 PR 05-FEB-1997; 97US-0795430.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 PA (UYHE-) UNIV HELSINKI LICENSING LTD.
 XX
 PI Alitalo K, Joukov V;
 XX
 DR WPI; 1998-437470/37.
 XX
 PR New isolated vascular endothelial growth factor polypeptide(s) -
 PR used to develop products for treating, e.g. cancers, inflammation,
 PR oedema, granulocytopenia or for wound healing or tissue
 PT transplantation
 PT
 PS Example 35; Page 143-145; 177pp; English.
 XX
 PS The vascular endothelial growth factor C (VEGF-C) polypeptides have
 CC activities affecting growth and migration of vascular endothelial cells,
 CC promoting growth of lymphatic endothelial cells and lymphatic vessels,
 CC increasing vascular permeability, and affecting myelopoiesis. The
 CC products can be used for stimulating angiogenesis, for inhibiting
 CC angiogenesis, for stimulating lymphangiogenesis, treatment or prevention
 CC of inflammation, oedema, elephantiasis, or Milroy's disease. They can
 CC also be used to modulate myelopoiesis, e.g. treating granulocytopenia.
 CC They can also be used for modulating the growth of endothelial cells.
 CC They can also be used to stimulate lymphocyte production and maturation,
 CC and to promote or inhibit trafficking of leucocytes between tissues and
 CC lymphatic vessels or to affect migration in and out of the thymus.
 XX
 SQ Sequence 419 AA;

Query Match 100.0%; Score 1218; DB 19; Length 419;
 Best Local Similarity 100.0%; Pred. No. 6e-90;
 Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CMSKLDVYQVHSIIIRSLPATLPQCAANKTCPTNYMNNHICRLAEDMFSSDAD 60
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 211 cmskldvyqvnslirslpatlpqcaanktcptnywmnhicrlaagedtmfssdagd 270
 QY 61 DSTGFHDICGNKELDEETCCVCVCRAGLRPASCGPHKELDRNSCQVCVCKNLFPSQCA 120
 Db |||||||fhdlcgnpkeldetccgcvcraqlrpascgphkeldrnsqcvcknlfpsqca 330
 QY 121 NREPDNTCCQVCKRTCPNPNPLNPGKACECTESPQCKLKGKFFHQTGSCYRRPCTN 180
 Db |||||||ntccqvcckrtcpnpglnpgkacectespqckllgkffhqtgscyrpctn 390
 QY 331 nrefdntccqvcckrtcpnpglnpgkacectespqckllgkffhqtgscyrpctn 390
 Db
 QY 181 ROKACEPGFSYSEVOCRCVPSYWKRPOMS 209
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 391 rkaacepgfsyseevocrvpsywkripms 419

RESULT 11
 AAW75740
 ID AAW75740 standard; Protein; 419 AA.

AC AAW75740;
 XX
 DT 20-NOV-1998 (first entry)
 XX
 DE Human vascular endothelial growth factor C protein.
 XX
 KW Flt4; vascular endothelial growth factor C; vascular endothelial cell;
 KW lymphatic endothelial cell; myelopoiesis; angiogenesis; inflammation;
 KW lymphangiogenesis; oedema; elephantiasis; Milroy's disease.
 XX
 OS Homo sapiens.
 XX
 PN WO9833917-A1.
 XX

PD 06-AUG-1998.
 XX
 PF 02-FEB-1998; 98WO-US01973.
 XX
 PR 05-FEB-1997; 97US-0795430.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 PA (UYHE-) UNIV HELSINKI LICENSING LTD.
 XX
 PI Alitalo K, Joukov V;
 XX
 DR WPI; 1998-437470/37.
 DR N-PSDB; AAV52576.
 XX
 PR New isolated vascular endothelial growth factor polypeptide(s) -
 PR used to develop products for treating, e.g. cancers, inflammation,
 PR oedema, granulocytopenia or for wound healing or tissue
 PT transplantation
 PT
 PS Claim 1; Page 112-115; 177pp; English.
 XX
 PS The vascular endothelial growth factor C (VEGF-C) polypeptides have
 CC activities affecting growth and migration of vascular endothelial cells,
 CC promoting growth of lymphatic endothelial cells and lymphatic vessels,
 CC increasing vascular permeability, and affecting myelopoiesis. The
 CC products can be used for stimulating angiogenesis, for inhibiting
 CC angiogenesis, for stimulating lymphangiogenesis, treatment or prevention
 CC of inflammation, oedema, elephantiasis, or Milroy's disease. They can
 CC also be used to modulate myelopoiesis, e.g. treating granulocytopenia.
 CC They can also be used for modulating the growth of endothelial cells.
 CC They can also be used to stimulate lymphocyte production and maturation,
 CC and to promote or inhibit trafficking of leucocytes between tissues and
 CC lymphatic vessels or to affect migration in and out of the thymus.
 XX
 SQ Sequence 419 AA;

Query Match 100.0%; Score 1218; DB 19; Length 419;
 Best Local Similarity 100.0%; Pred. No. 6e-90;
 Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CMSKLDVYQVHSIIIRSLPATLPQCAANKTCPTNYMNNHICRLAEDMFSSDAD 60
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 211 cmskldvyqvnslirslpatlpqcaanktcptnywmnhicrlaagedtmfssdagd 270
 QY 61 DSTGFHDICGNKELDEETCCVCVCRAGLRPASCGPHKELDRNSCQVCVCKNLFPSQCA 120
 Db |||||||fhdlcgnpkeldetccgcvcraqlrpascgphkeldrnsqcvcknlfpsqca 330
 QY 121 NREPDNTCCQVCKRTCPNPNPLNPGKACECTESPQCKLKGKFFHQTGSCYRRPCTN 180
 Db |||||||ntccqvcckrtcpnpglnpgkacectespqckllgkffhqtgscyrpctn 390
 QY 331 nrefdntccqvcckrtcpnpglnpgkacectespqckllgkffhqtgscyrpctn 390
 Db
 QY 181 ROKACEPGFSYSEVOCRCVPSYWKRPOMS 209
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 391 rkaacepgfsyseevocrvpsywkripms 419

RESULT 12
 AAW86203
 ID AAW86203 standard; Protein; 419 AA.

AC AAW86203;
 XX
 DT 16-FEB-1999 (first entry)
 XX
 DE Human vascular endothelial growth factor (VEGF)-C sequence.
 XX
 KW VEGF; VEGF; vascular endothelial growth factor; VEGF-related protein;
 KW recombinant; truncated; gene therapy; angiogenesis; cardiac ischemia;
 KW coronary; collateral vessel development; cell growth; migration; heart;
 KW lower limb ischemia; stroke; peripheral vascular disease; intestine;
 KW wound healing; skin; vascular permeability.

```

XX Homo sapiens.
OS
XX MO9849300-A2.
PN
XX
XX 05-NOV-1998.
PD
XX
XX 20-APR-1998; 98WO-US07801.
PF
XX
XX 25-APR-1997; 97US-0842984.
PR
XX
XX (COLL-) COLLATERAL THERAPEUTICS.
PA
XX
XX Bohlen P;
PI
XX
XX WPI: 1999-009426/01.
DR
XX
XX New truncated vascular endothelial growth factor-related protein
PT subunits - lack part of the N-terminal sequence, used to stimulate
PT angiogenesis, e.g. for treating heart disease and ischemia
PS
XX
XX Disclosure: Fig 1; 113pp; English.
XX
XX This represents the amino acid sequence of human vascular endothelial
XX growth factor (VEGF)-C protein. The invention provides truncated VRP
XX (VEGF-related protein) subunits that have at least one amino acid
XX N-terminal to the first Cys of the core sequence deleted. Host cells
XX transformed or transfected with expression vectors containing nucleic
XX acids encoding the truncated VRP subunits are used to produce the
XX truncated proteins recombinantly. The truncated VRP subunits, optionally
XX expressed from gene therapy vectors, have in vivo and in vitro angiogenic
XX activity and are used to stimulate angiogenesis, particularly coronary
XX collateral vessel development in cases of cardiac ischemia; to stimulate
XX endothelial cell growth and migration in vitro; to treat heart disease;
XX to treat ischemia (e.g. cardiac, chronic coronary or chronic lower limb
XX ischemia; stroke and peripheral vascular disease); to promote healing of
XX wounds (of skin or intestines), and to increase vascular permeability.
XX
XX Sequence 419 AA:
SQ
Query Match 100.0%; Score 1218; DB 20; Length 419;
Best Local Similarity 100.0%; Pred. No. 6e-90;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CMSKIDVYRVQVSIIRSRSLPATLPQCAANKTCPTNYMNMNHCRLAQEDMFSSDAGD 60
DB 211 CMSKIDVYRVQVSIIRSRSLPATLPQCAANKTCPTNYMNMNHCRLAQEDMFSSDAGD 270
QY 61 DSTDFHDIICGNKELDEETCCVCVRAGLRPASCGPHKELDRNSCCVCVKNKLFPSQCA 120
DB 271 dstdghdiciqgnkeldetccvcvraglrpsacgphkeldrnsccvcvknklfpsqca 330
QY 121 NREPDENTCCVCVKRTCPRNQPLNPGKACACECTESPQCKLKGKFFHNOTGSCYRRPCTN 180
DB 331 nreidentccvcvkrtcprnqplnpgkacacetespqckllkgkffhnotgscyrpctn 390
QY 181 ROKACEPGFSYSSEVCRCPVSYWKRPM 209
DB 391 rqacepgfysseevcrvpsywkrpms 419

```

RESULT 13

AAB10648 standard: Protein: 419 AA.

AAB10648;

19-JAN-2001 (first entry)

Human VEGC protein.

VEGF-X; vascular endothelial growth factor; human; vulnery; cytostatic;

```

KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
KW venous sore; diabetic ulcer; burns; skin graft growth; VEGC.
XX
XX Homo sapiens.
OS
XX
XX MO200037641-A2.
PN
XX
XX 29-JUN-2000.
PD
XX
XX 21-DEC-1999; 99WO-US30503.
PF
XX
XX 22-DEC-1998; 98GB-0028377.
PR 18-MAR-1999; 99US-0124967.
PR 08-NOV-1999; 99US-0164131.
XX
XX (JANC ) JANSSEN PHARM NV.
PA
XX
XX Gordon RD, Sprengel JJ, Yon JR, Dijkmans JH, Gostlewska A;
PI Dhanraj SN, Xu J;
XX
XX WPI: 2000-442669/38.
DR
XX
XX New vascular endothelial growth factor protein, useful for treating or
XX preventing diseases associated with inappropriate angiogenesis activity
XX such as cancer, rheumatoid arthritis, psoriasis and wounds -
XX
XX Disclosure: Fig 11; 127pp; English.
XX
XX This invention describes a novel vascular endothelial growth factor-X
XX (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
XX vulnery cytostatic, antirheumatic, antiarthritic, antipsoriatic and
XX antidiabetic activity, and acts as an angiogenesis and vascularization
XX regulator. An antisense molecule of the invention is useful for treating
XX or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
XX retinopathy by inhibiting angiogenic activity or inappropriate
XX vascularization including formation and proliferation of new blood
XX vessels, growth and development of tissues, tissue regeneration and organ
XX and tissue repair in a subject. The products of the invention are useful
XX for preparing medicaments for treating wounds such as dermal ulcers,
XX pressure sores, venous sores, diabetic ulcers and burns and to promote
XX skin graft growth, tissue repair, proliferation of new blood vessels,
XX tissue regeneration and organ repair by promoting angiogenic activity or
XX vascularization. This sequence represents the human VEGC protein used
XX to illustrate the method of the invention.
XX
XX Sequence 419 AA:
SQ
Query Match 100.0%; Score 1218; DB 21; Length 419;
Best Local Similarity 100.0%; Pred. No. 6e-90;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CMSKIDVYRVQVSIIRSRSLPATLPQCAANKTCPTNYMNMNHCRLAQEDMFSSDAGD 60
DB 211 CMSKIDVYRVQVSIIRSRSLPATLPQCAANKTCPTNYMNMNHCRLAQEDMFSSDAGD 270
QY 61 DSTDFHDIICGNKELDEETCCVCVRAGLRPASCGPHKELDRNSCCVCVKNKLFPSQCA 120
DB 271 dstdghdiciqgnkeldetccvcvraglrpsacgphkeldrnsccvcvknklfpsqca 330
QY 121 NREPDENTCCVCVKRTCPRNQPLNPGKACACECTESPQCKLKGKFFHNOTGSCYRRPCTN 180
DB 331 nreidentccvcvkrtcprnqplnpgkacacetespqckllkgkffhnotgscyrpctn 390
QY 181 ROKACEPGFSYSSEVCRCPVSYWKRPM 209
DB 391 rqacepgfysseevcrvpsywkrpms 419

```

RESULT 14

ID	AA#29048	standard: Protein; 419 AA.
XX	AA#29048;	
XX	31-JAN-2001	(first entry)
XX		Human VEGF-C protein sequence.
XX		Human: Flt4; fms-like tyrosine kinase 4; lymphoedema;
XX		vascular endothelial growth factor receptor 3; VEGFR-3;
XX		Milroy-Nonne syndrome; lymphoedema praecox; VEGF-C;
XX		vascular endothelial growth factor C.
OS	Homo sapiens.	
PN	WO200058511-A1.	
PD	05-OCT-2000.	
PE	26-MAR-1999;	99WO-US06133.
PR	26-MAR-1999;	99WO-US06133.
XX		(LUDWIG INST CANCER RES.
XX		(UYHE-) UNIV HELSINKI LICENSING LTD OY.
XX		(UYPI-) UNIV PITTSBURGH.
PI	Ferrell RE, Altalo K, Finegold DN, Karkkainen M;	
DR	WPI; 2000-679298/66.	
DR	N-PSDB; AAC62406.	
XX		Screening a human subject for increased risk of developing a lymphatic
PT	disorder, comprises assaying a nucleic acid to determine a mutation	
PT	altering the sequence of a vascular endothelial growth factor	
PT	receptor-3 -	
PS	Disclosure; Page 60-61; 76pp; English.	
XX		The present sequence is the protein sequence for the human vascular
CC		endothelial growth factor C (VEGF-C). It was used to demonstrate the
CC		methods of the invention, which involve the screening of individuals to
CC		determine which vascular endothelial growth factor receptor 3 (VEGFR-3,
CC		also known as Flt4 or fms-like tyrosine kinase 4) alleles they possess
CC		and thus their likelihood of developing hereditary lymphoedema.
CC		Conditions associated with lymphoedema include Milroy-Nonne syndrome,
CC		which is early onset lymphoedema and lymphoedema praecox, which is late
CC		onset.
SO	Sequence 419 AA;	
Query Match	100.0%; Score 1218; DB 21; Length 419;	
Best Local Similarity	100.0%; Pred. No. 6e-90;	
Matches 209; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
1	CMSTKIDYVROYVSIIRRSIPATLPQCOQANKTCPTNYMNNHICICLAQEDFMFSSDAGD	60
211	CMSTKIDYVRYVSIIRRSIPALIPQCGAANKTCPTNYMNNHICICLAQEDFMFSSDAGD	270
61	DSTDGFHICIGCNKELDEETCCQVCVRAGLRPASGCPHRELDSNCCQVCKNKLPSQCA	120
271	dstdghfhtcgnknldeetccqvcvragslrasgcphehldsnscqvcxknklpsqca	330
121	NREEDENTCQCVCARCTCPRNPDLNPGKCAECTESPQKCLLKGKKEHMQTSCYRRPCTN	180
331	nreidentcgcckrtcpnqplnpqkacectespqkcllkgkkfhmqtcscyrpctn	390
181	RKACEPGEFSYSEECVRCVPSYWRPQMS	209
391	rkacepgfisyseevrcvpsyrwkrpqms	219

	RESULT	15
AA70749	ID	AAY70749 standard; Protein; 419 AA.
XX	AC	
XX	DT	17-AUG-2000 (first entry)
XX	DE	Human prepro-vascular endothelial growth factor C.
XX	KW	Human; receptor tyrosine kinase; RTK; Flt4; fms-like tyrosine kinase 4;
XV	KW	VEGFR-3; vascular endothelial growth factor receptor-3; chromosome 5q35;
KX	KW	cystostatic; tumour imaging; anti-tumour therapy; treatment; diagnosis;
KY	KW	neoplastic disease; lymphoma; carcinoma; breast; squamous cell; melanoma;
KZ	KW	sarcoma; malignancy; VEGF-C; vascular endothelial growth factor C.
OS	Homo sapiens.	
XX	Key	Location/Qualifiers
FH	Peptide	/label= Signal-peptide
FT	Peptide	32..103
FT	Peptide	/label= N-terminal_peptide
FT	Protein	/note= "cleavage of this peptide from partially processed VEGF-C produces a fully processed mature form of VEGF-C of 21-23 kD which has high affinity to VEGFR-2"
FT	Peptide	104..227
FT	Peptide	/label= Mature_VEGF-C
FT	Peptide	228..419
FT	Peptide	/label= C-terminal_peptide
FT	Region	/note= "Has a pattern of spaced cysteine residues reminiscent of a Balbiani ring 3 protein (BR3P) sequence; cleavage of signal peptide and the C-terminal peptide produces a partially processed form of VEGF-C of about 29 kD which has high affinity to Flt4 (VEGFR-3)" 113..213
FT	Binding-site	/note= "binds and stimulates VEGF-C receptors; Cys at position 156 is essential for VEGFR-2 binding and at 165 is essential for VEGFR-2 and VEGFR-3 binding" 131..211
FT	Region	/note= "Important for VEGF-C activity"
FN	MO200021560-Al.	
PJ	PD	20-APR-2000.
PP	XX	08-OCT-1999; 99WO-US23525.
PR	XX	09-OCT-1998; 98US-0169079.
PA	(LUDWI-) LUDWIG INST CANCER RES. (UYHE-) UNIV HELSINKI LICENSING LTD OY.	
PI	Allitalo K, Kaipainen A, Vaitola R, Jussila E;	
DQ	WP1; 2000-317850/27.	
PT	Treating neoplastic diseases such as lymphoma, carcinomas, melanomas and sarcomas, involves administering a compound capable of inhibiting binding of ligand proteins to fms-like tyrosine kinase-4 receptor -	
PS	Example 15-17; Page 140-142; 148pp; English.	
XX	The patent discloses a method to treat neoplastic disease characterised by expression of fms-like tyrosine kinase 4 (Flt4) receptor (also referred as vascular endothelial growth factor receptor-3, VEGFR-3) in endothelial cells of blood vessels adjacent to malignant neoplasm. The method involves administering a compound that inhibits binding of a ligand to Flt4 thereby inhibiting Flt4 mediated proliferation of vascular endothelial cells. The compound is useful for treating neoplastic disease such as breast carcinomas, squamous cell carcinomas, lymphomas, melanomas and sarcomas. Flt4 receptor tyrosine kinase binding compounds can be used	

CC for manufacturing medicament useful for diagnostic screening, imaging and
CC treatment of malignancies characterised by Flt4-expressing blood cells.
CC The Flt4 gene maps to chromosomal region 5q35 and is expressed as 5.8 kb
CC and 4.5 kb mRNAs which differ in their 3' sequences and are
CC differentially expressed in HEL and DAMI cell lines. Flt4
CC belongs to a subfamily of class III receptor tyrosine kinases (RTKs).
CC It is used as a target for tumour imaging and anti-tumour therapy.
CC The present sequence is a human prepro-vascular endothelial growth
CC factor C (VEGF-C), a specific example of Flt4 binding compound.
XX
SO Sequence 419 AA:

Query Match	100.0%	Score 1218;	DB 21;	Length 419;
Best Local Similarity	100.0%	Pred. No. 6e-90;		
Matches 209; Conservative	0;	Mismatches	0;	Gaps 0

QY	1	CSMSKIDVYQVHSTIRSLPTLEQQCAANKTCYTNMNNHILRCILAOEDPFMSDAG	60
Db	211	cmkskldvyqghstirslptlcpcqaanktcptnywmnhlrcrlaagedfmfsdagd	270
QY	61	DSTDGFHDICGPNKELDEETCCQCVCRAGLRPASCGPKHELDNRMSQCVCNKKLFPSCGA	120
Db	271	dstdghfhdicgpnkeldetccqcvcraglrpascgpkhieldnrmscqvcvcknlfpscgga	330
QY	121	NREDEWNTCCQVCCKRTCPRNQPLNPGKCAEBCTFSPQCKLLGKKTFHMQTSCSYRRPCTN	180
Db	331	nreidentccqvcckrtcptprnlmpgkcaectespqckllgkklfbhgqtcscyrpctn	390
QY	181	RQKACEPGFSYSEWCRCVPSYMKRPQMS	209
Db	391	rqkacepgfisyseewrcvpwymrkpqms	419

```
Search completed: October 17, 2001, 14:46:00
Job time: 126 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:47:09 ; Search time 62.93 seconds
(without alignments)
68.384 Million cell updates/sec

Title: US-09-427-657-2_COPY_211_419

Perfect score: 1218

Sequence: 1 CMSKLDVRYOVHSIIRSLP.....SYSEYRCVPSYKRPQMS 209

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 197339 segs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents-AA: *
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PTOUS_COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1218	100.0	350	US-08-510-133A-33	Sequence 33, Appl
2	1218	100.0	350	US-08-585-895-33	Sequence 33, Appl
3	1218	100.0	419	US-08-795-430-8	Sequence 8, Appl
4	1214	100.0	419	US-08-510-133A-35	Sequence 35, Appl
5	1214	99.7	350	US-08-999-811-4	Sequence 4, Appl
6	1214	99.7	350	US-08-824-996-2	Sequence 2, Appl
7	1214	99.7	419	US-09-042-105-4	Sequence 4, Appl
8	1214	99.7	419	US-08-999-811-2	Sequence 2, Appl
9	1214	99.7	419	US-09-042-105-2	Sequence 2, Appl
10	1214	99.7	419	US-09-042-105-18	Sequence 18, Appl
11	1214	99.7	419	PCT-US96-09001-2	Sequence 2, Appl
12	1086	89.2	415	US-08-795-430-11	Sequence 11, Appl
13	927	76.1	418	US-08-795-430-13	Sequence 13, Appl
14	253.5	20.8	325	US-08-915-795-3	Sequence 3, Appl
15	253.5	20.8	354	US-08-915-795-5	Sequence 5, Appl
16	241.5	19.8	358	US-08-915-795-8	Sequence 8, Appl
17	205.5	16.9	321	US-08-915-795-9	Sequence 9, Appl
18	151	12.4	1810	PCT-US95-11684-4	Sequence 4, Appl
19	143.5	11.8	2813	US-08-896-449A-2	Sequence 2, Appl
20	143.5	11.8	2813	US-09-132-652-2	Sequence 2, Appl
21	139.5	11.5	2471	US-08-185-432-16	Sequence 16, Appl
22	139.5	11.5	2471	US-08-083-590A-19	Sequence 19, Appl
23	139.5	11.5	2471	US-08-537-384-19	Sequence 19, Appl
24	137.5	11.3	2523	US-08-185-432-18	Sequence 18, Appl
25	134	11.0	1219	US-08-862-046-5	Sequence 5, Appl
26	133	10.9	642	US-08-872-855-10	Sequence 10, Appl
27	130	10.7	1010	US-08-882-046-7	Sequence 7, Appl

28	130	10.7	1218	2	US-08-400-159-6	Sequence 6, Appl
29	130	10.7	1218	3	US-08-611-729A-6	Sequence 6, Appl
30	130	10.7	1218	4	US-08-882-046-2	Sequence 2, Appl
31	130	10.7	2199	5	PCT-US95-11684-2	Sequence 2, Appl
32	128	10.5	2703	1	US-08-185-432-19	Sequence 19, Appl
33	127.5	10.5	2556	1	US-08-083-590A-20	Sequence 20, Appl
34	127.5	10.5	2556	1	US-08-532-384-20	Sequence 20, Appl
35	125.5	10.3	374	2	US-08-820-170A-25	Sequence 25, Appl
36	125.5	10.3	374	3	US-09-055-699-25	Sequence 25, Appl
37	125.5	10.3	374	4	US-09-273-565-25	Sequence 25, Appl
38	125.5	10.3	374	5	PCT-US95-06385-2	Sequence 2, Appl
39	124.5	10.2	2556	1	US-08-185-432-17	Sequence 17, Appl
40	124	10.2	314	2	US-08-460-309-19	Sequence 19, Appl
41	124	10.2	314	2	US-08-125-077-19	Sequence 19, Appl
42	123	10.1	2050	2	US-08-347-594A-2	Sequence 2, Appl
43	122	10.0	1193	2	US-08-400-159-10	Sequence 10, Appl
44	122	10.0	1193	3	US-08-611-729A-10	Sequence 10, Appl
45	121	9.9	1358	1	US-08-404-665-4	Sequence 4, Appl

ALIGNMENTS

```
RESULT 1
US-08-510-133A-33
: Sequence 33, Application US/08510133A
: Patent No. 6221839
:
: GENERAL INFORMATION:
:   APPLICANT: Alltalo, Kari
:             Joukov, Vladimir
: TITLE OF INVENTION: Receptor Ligand
: NUMBER OF SEQUENCES: 35
: CORRESPONDENCE ADDRESS:
:   ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
:   STREET: 6300 Sears Tower, 233 South Wacker Drive
:   CITY: Chicago
:   STATE: Illinois
:   COUNTRY: United States of America
:   ZIP: 60606-6402
: COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk
:   OPERATING SYSTEM: PC-DOS/MS-DOS
:   SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/08/510,133A
:   FILING DATE: 01-Aug-1995
:   CLASSIFICATION: <Unknown>
:   ATTORNEY/AGENT INFORMATION:
:     NAME: Gass, David A.
:     REGISTRATION NUMBER: 38,153
:     REFERENCE/DOCKET NUMBER: 28113/32863
:   TELECOMMUNICATION INFORMATION:
:     TELEPHONE: 312/474-6300
:     TELEFAX: 312/474-0448
:     TELEX: 25-3856
: INFORMATION FOR SEQ ID NO: 33:
:   SEQUENCE CHARACTERISTICS:
:     LENGTH: 350 amino acids
:     TYPE: amino acid
:     TOPOLOGY: linear
:   MOLECULE TYPE: protein
:   SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-08-510-133A-33

Query Match      100.0%; Score 1218; DB 4; Length 350;
Best Local Similarity 100.0%; Pred. No. 1.4e-104;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CMSKLDVRYOVHSIIRSLPATLPCQCAANKTCPTNYMNNHICRCLAQEDFMFSSDAG 60
Db      142 CMSKLDVRYOVHSIIRSLPATLPCQCAANKTCPTNYMNNHICRCLAQEDFMFSSDAG 201
```


QY	61	DSTGDFHICPNKNELBEEETQCVRAGLRPASGPKHKLDRNSCCQVCNKXLPSPCGA	120
Db	271	DSTGDFHICPNKNELBEEETQCVRAGLRPASGPKHKLDRNSCCQVCNKXLPSPCGA	330
QY	121	NREDFENTCQCVRKTCPRNOLPLNPGKACACECTESPCKLTKGKFFHQTSCYRRCCTN	180
Db	331	NREDFENTCQCVRKTCPRNOLPLNPGKACACECTESPCKLTKGKFFHQTSCYRRCCTN	390
QY	181	RQKACEPGFSYSEFVRCRVPSTYMKRPOMS	209
Db	391	RQKACEPGFSYSEFVRCRVPSTYMKRPOMS	419

```

1      RESULT 4
2      US-08-510-133A-35
3      Sequence 35: Application US/08510133A
4      Patent No. 6221839
5      GENERAL INFORMATION:
6      APPLICANT: Altalo, Kari
7      Joukov, Vladimir
8      TITLE OF INVENTION: Receptor Ligand
9      NUMBER OF SEQUENCES: 35
10     CORRESPONDENCE ADDRESS:
11     ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
12     STREET: 6300 Sears Tower, 233 South Wacker Drive
13     CITY: Chicago
14     STATE: Illinois
15     COUNTRY: United States of America
16     ZIP: 60606-6402
17
18     COMPUTER READABLE FORM:
19     MEDIUM TYPE: floppy disk
20     COMPUTER: IBM PC compatible
21     OPERATING SYSTEM: PC-DOS/MS-DOS
22     SOFTWARE: Patent In Release #1.0, Version #1.25
23
24     CURRENT APPLICATION DATA:
25     APPLICATION NUMBER: US/08/510.133A
26     FILING DATE: 01-Aug-1995
27     CLASSIFICATION: <Unknown>
28     ATTORNEY/AGENT INFORMATION:
29     NAME: Gass, David A.
30     REGISTRATION NUMBER: 38,153
31     REFERENCE/DOCKET NUMBER: 28113/32863
32     TELECOMMUNICATION INFORMATION:
33     TELEPHONE: 312/474-6300
34     TELEFAX: 312/474-0448
35     TELEX: 25-3856
36
37     INFORMATION FOR SEQ ID NO: 35:
38     SEQUENCE CHARACTERISTICS:
39     LENGTH: 419 amino acids
40     TYPE: amino acid
41     TOPOLOGY: linear
42     MOLECULE TYPE: protein
43     SEQUENCE DESCRIPTION: SEQ ID NO: 35:
44     US-08-510-133A-35

```

	Query Match	Best Local Similarity	100.0%;	Score 1218;	DB 4;	Length 419;
	Matches	Conservative	0;	Pred. No. 1.7e-104;	Mismatches 0;	Indels 0; Gaps
QY	1	CSKSLDLYRV	ROVHSII	RRSLP	POCCAA	NKTCPTNMYNNNHICRLCLAEDEMFSSDAG 60
Db	211	CSKSLDLYRV	ROVHSII	RRSLP	POCCAA	NKTCPTNMYNNNHICRLCLAEDEMFSSDAG 270
QY	61	DSTGPFHDICG	PNKELEDE	EOCCCR	AGLTPAS	CGPHEKELDRNSCCQVCVCKNKKTFPSQCA 120
Db	271	DSTGPFHDICG	PNKELEDE	EOCCCR	AGLTPAS	CGPHEKELDRNSCCQVCVCKNKKTFPSQCA 330
QY	121	NREDEDTNCQ	CVCKRTCP	PRN	QPLNPG	CAACECTESPQCLLKSKKFFHHQSCSYRRPCTN 180
Db	331	NREDEDTNCQ	CVCKRTCP	PRN	QPLNPG	CAACECTESPQCLLKSKKFFHHQSCSYRRPCTN 390
QY	181	RQKACEFGFST	SEVCEVC	VCVSY	NKRPDM	S 209

```

Db      391  RÖKACEPGFSYSEEVCRVCPSYWKRPÖMS  419

```

```

RESULT 5
US-08-999-811-4
: Sequence 4, Application US/08999811
: Patent No. 5932540
:
: GENERAL INFORMATION:
: APPLICANT: HU, JING-SHAN
: APPLICANT: ROSEN, CRAIG A.
: APPLICANT: CAO, LIANG
: TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
: NUMBER OF SEQUENCES: 15
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX
: STREET: 1100 NEW YORK AVENUE
: CITY: WASHINGTON
: STATE: DC
: COUNTRY: USA
: ZIP: 20005
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/999,811
: FILING DATE: HEREWITH
: CLASSIFICATION:
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/207,550
: FILING DATE: 8-MAR-1994
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/465,968
: FILING DATE: 06-JUN-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: MAROWICZ, KAREN R.
: REGISTRATION NUMBER: 36,351
: REFERENCE/DOCKET NUMBER: 1488.1000004
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)371-2600
: TELEFAX: (202)371-2540
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 350 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
: US-08-999-811-4

```

	Query Match	99.7%;	Score 1214;	DB 2;	Length 350;
	Best Local Similarity	99.5%;	Pred. No. 3,36-104;		
	Matches 208;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CMSKLDVYROYHSIIRSLPATLPPCQANKTCPTNYMNNHICICLQAOEDFMSSDAGD	60		
Db	142	CMSKLDVYROYHSIIRSLPATLPPCQANKTCPTNYMNNHICICLQAOEDFMSSDAGD	201		
QY	61	DSTGPFHICGPNKELEBETCCVCRAALRPASCGPHKEIDRNSCCGCKNKLFPSSCGA	120		
Db	202	DSTGPFHICGPNKELEBETCCVCRAALRPASCGPHKEIDRNSCCGCKNKLFPSSCGA	261		
QY	121	NREFDENTCCGCVKRCRTCPRNQPLNPGKACACBCTESPQCLLKGGKFFHQTCSYRRPCTN	180		
Db	262	NREFDENTCCGCVKRCRTCPRNQPLNPGKACACBCTESPQCLLKGGKFFHQTCSYRRPCTN	321		
QY	181	RQKACBPGFSYSEVYGCRCVPSTMYKRPQMS	209		
Db	322	RQKACBPGFSYSEVYGCRCVPSTMYKRPQMS	350		

RESULT 6
US-08-824-996-2
Sequence 2, Application US/08824996B
Patent No. 5935820
GENERAL INFORMATION:
APPLICANT: Hu, Jing-Shan
APPLICANT: Rosen, Craig A.
APPLICANT: Cao, Liang
TITLE OF INVENTION: Polynucleotides Encoding Vascular Endothelial Growth
TITLE OF INVENTION: Factor 2
FILE REFERENCE: PFI12D1
CURRENT APPLICATION NUMBER: US/08/824,996B
CURRENT FILING DATE: 1997-03-27
EARLIER APPLICATION NUMBER: 08/207,550
EARLIER FILING DATE: 1994-03-08
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 350
TYPE: PRT
ORGANISM: Homo sapiens
US-08-824-996-2

Query Match 99.7%; Score 1214; DB 2; Length 350;
Best Local Similarity 99.5%; Pred. No. 3.3e-104;
Matches 208; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CMSKLDVYQVHSIIIRSLPATLPQQAANKTCPTNYMMNNHICRCLAQEDFMFSSDAGD 60
DB 142 CMSKLDVYQVHSIIIRSLPATLPQQAANKTCPTNYMMNNHICRCLAQEDFMFSSDAGD 201
QY 61 DSTDGFHDICGNKELDEETCCQCVCRAGLRPASCGPHKELDRNSCQCVCKNKLFPSCGCA 120
DB 202 DSTDGFHDICGNKELDEETCCQCVCRAGLRPASCGPHKELDRNSCQCVCKNKLFPSCGCA 261
QY 121 NREFDNTCCQCVCKRCPNPNPLNPGKCAECTESPQKCLLGKKFHHOTCSCYRRPCTN 180
DB 262 NREFDNTCCQCVCKRCPNPNPLNPGKCAECTESPQKCLLGKKFHHOTCSCYRRPCTN 321
QY 181 ROKACEPGFSYSEVCRCPVSYWKRQMS 209
DB 322 ROKACEPGFSYSEVCRCPVSYWKRQMS 350

RESULT 7
US-09-042-105-4
Sequence 4, Application US/09042105
Patent No. 6040157
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,105
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550

FILING DATE: 8-MAR-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TO BE ASSIGNED
FILING DATE: 24-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: ERIC K. STEFFE
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-042-105-4

Query Match 99.7%; Score 1214; DB 3; Length 350;
Best Local Similarity 99.5%; Pred. No. 3.3e-104;
Matches 208; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CMSKLDVYQVHSIIIRSLPATLPQQAANKTCPTNYMMNNHICRCLAQEDFMFSSDAGD 60
DB 142 CMSKLDVYQVHSIIIRSLPATLPQQAANKTCPTNYMMNNHICRCLAQEDFMFSSDAGD 201
QY 61 DSTDGFHDICGNKELDEETCCQCVCRAGLRPASCGPHKELDRNSCQCVCKNKLFPSCGCA 120
DB 202 DSTDGFHDICGNKELDEETCCQCVCRAGLRPASCGPHKELDRNSCQCVCKNKLFPSCGCA 261
QY 121 NREFDNTCCQCVCKRCPNPNPLNPGKCAECTESPQKCLLGKKFHHOTCSCYRRPCTN 180
DB 262 NREFDNTCCQCVCKRCPNPNPLNPGKCAECTESPQKCLLGKKFHHOTCSCYRRPCTN 321
QY 181 ROKACEPGFSYSEVCRCPVSYWKRQMS 209
DB 322 ROKACEPGFSYSEVCRCPVSYWKRQMS 350

RESULT 8
US-08-999-811-2
Sequence 2, Application US/08999811
Patent No. 5932540
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/999,811
FILING DATE: HERewith
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550
FILING DATE: 8-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: MARKOWICZ, KAREN R.
REGISTRATION NUMBER: 36,351
REFERENCE/DOCKET NUMBER: 1488.1000004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-999-811-2

Query Match 99.7%; Score 1214; DB 2; Length 419;
Best Local Similarity 99.5%; Pred. No. 4e-104;
Matches 208; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CMSKLDVYRQVHSIIRSLPATLPCCQANKTCPTNYMNNHICRLAOEDMFSSDAGD 60
DB 211 CMSKLDVYRQVHSIIRSLPATLPCCQANKTCPTNYMNNHICRLAOEDMFSSDAGD 270
QY 61 DSTDFHDCGNKELDEETCCQVCAGLRPASCGPHKELDRNSCQVCCKNLFPSQCGA 120
DB 271 DSTDFHDCGNKELDEETCCQVCAGLRPASCGPHKELDRNSCQVCCKNLFPSQCGA 330
QY 121 NNEPENTCQCCYCKRTCPNPNPLNPGKACCECTESPQCKLKGKFFHQTCSYRRPCTN 180
DB 331 NNEPENTCQCCYCKRTCPNPNPLNPGKACCECTESPQCKLKGKFFHQTCSYRRPCTN 390
QY 181 ROKACEPGFSYSEVCRCPVSYWRRPOMS 209
DB 391 ROKACEPGFSYSEVCRCPVSYWRRPOMS 419

RESULT 9
US-09-042-105-2
Sequence 2, Application US/09042105
Patent No. 6040157
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,105
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550
FILING DATE: 8-MAR-1994
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TO BE ASSIGNED
FILING DATE: 24-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: ERIC K. STEFFE
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-042-105-2

Query Match 99.7%; Score 1214; DB 3; Length 419;
Best Local Similarity 99.5%; Pred. No. 4e-104;
Matches 208; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CMSKLDVYRQVHSIIRSLPATLPCCQANKTCPTNYMNNHICRLAOEDMFSSDAGD 60
DB 211 CMSKLDVYRQVHSIIRSLPATLPCCQANKTCPTNYMNNHICRLAOEDMFSSDAGD 270
QY 61 DSTDFHDCGNKELDEETCCQVCAGLRPASCGPHKELDRNSCQVCCKNLFPSQCGA 120
DB 271 DSTDFHDCGNKELDEETCCQVCAGLRPASCGPHKELDRNSCQVCCKNLFPSQCGA 330
QY 121 NNEPENTCQCCYCKRTCPNPNPLNPGKACCECTESPQCKLKGKFFHQTCSYRRPCTN 180
DB 331 NNEPENTCQCCYCKRTCPNPNPLNPGKACCECTESPQCKLKGKFFHQTCSYRRPCTN 390
QY 181 ROKACEPGFSYSEVCRCPVSYWRRPOMS 209
DB 391 ROKACEPGFSYSEVCRCPVSYWRRPOMS 419

RESULT 10
US-09-042-105-18
Sequence 18, Application US/09042105
Patent No. 6040157
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,105
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550

FILING DATE: 8-MAR-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TO BE ASSIGNED
FILING DATE: 24-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: ERIC K. STEFFE
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2340
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-042-105-18

Query Match 99.7%; Score 1214; DB 3; Length 419;
Best Local Similarity 99.5%; Pred. No. 4e-104;
Matches 208; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CMSKLDVYRQVHSIIRSLPATLPCCQAANKTCPTNYMMNNHICRCLAOEDFMFSSDAG 60
DB 211 CMSKLDVYRQVHSIIRSLPATLPCCQAANKTCPTNYMMNNHICRCLAOEDFMFSSDAG 270
QY 61 DSTDGFHDICGNKELDEETCCQVCVAGLRPASCGPHKELDRNSCCQVCVCKNKLFPSSQCA 120
DB 271 DSTDGFHDICGNKELDEETCCQVCVAGLRPASCGPHKELDRNSCCQVCVCKNKLFPSSQCA 330
QY 121 NREFDNTCCQCKKRTCPRNOPPLNPGKACCECTESPQCKLLGKKFHHQTCGYRRPCTN 180
DB 331 NREFDNTCCQCKKRTCPRNOPPLNPGKACCECTESPQCKLLGKKFHHQTCGYRRPCTN 390
QY 181 ROKACEPGFSYSEVCRCPYSYWRPOMS 209
DB 391 ROKACEPGFSYSEVCRCPYSYWRPOMS 419

RESULT 11

PCT-US96-09001-2
Sequence 2, Application PC/TUS9609001

GENERAL INFORMATION:
APPLICANT: HU, ET AL.
TITLE OF INVENTION: Human Vascular Endothelial Growth Factor 2
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: CAREELA BYRNE, BAIN, GILFILLAN,
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/09001
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/465,968
FILING DATE: 6 JUN 95
APPLICATION NUMBER: 08/207,550
FILING DATE: 8 MAR 1994
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-288
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
PCT-US96-09001-2

Query Match 99.7%; Score 1214; DB 5; Length 419;
Best Local Similarity 99.5%; Pred. No. 4e-104;
Matches 208; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CMSKLDVYRQVHSIIRSLPATLPCCQAANKTCPTNYMMNNHICRCLAOEDFMFSSDAG 60
DB 211 CMSKLDVYRQVHSIIRSLPATLPCCQAANKTCPTNYMMNNHICRCLAOEDFMFSSDAG 270
QY 61 DSTDGFHDICGNKELDEETCCQVCVAGLRPASCGPHKELDRNSCCQVCVCKNKLFPSSQCA 120
DB 271 DSTDGFHDICGNKELDEETCCQVCVAGLRPASCGPHKELDRNSCCQVCVCKNKLFPSSQCA 330
QY 121 NREFDNTCCQCKKRTCPRNOPPLNPGKACCECTESPQCKLLGKKFHHQTCGYRRPCTN 180
DB 331 NREFDNTCCQCKKRTCPRNOPPLNPGKACCECTESPQCKLLGKKFHHQTCGYRRPCTN 390
QY 181 ROKACEPGFSYSEVCRCPYSYWRPOMS 209
DB 391 ROKACEPGFSYSEVCRCPYSYWRPOMS 419

RESULT 12

US-08-795-430-11
Sequence 11, Application US/08795430

GENERAL INFORMATION:
APPLICANT: Alltalo, Kari
APPLICANT: Joukov, Vladimir
TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,430
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IT96/00427
FILING DATE: 01-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/671,573


```

1      PFLING DATE: 28-JUN-1996
2      PRIOR APPLICATION DATA:
3      APPLICATION NUMBER: 08/601,132
4      FILING DATE: 14-FEB-1996
5      PRIOR APPLICATION DATA:
6      APPLICATION NUMBER: 08/585,895
7      FILING DATE: 12-JAN-1996
8      PRIOR APPLICATION DATA:
9      APPLICATION NUMBER: 08/510,133
10     FILING DATE: 01-AUG-1995
11     PRIOR APPLICATION DATA:
12     APPLICATION NUMBER: 08/340,011
13     FILING DATE: 14-NOV-1994
14     ATTORNEY/AGENT INFORMATION:
15     NAME: Gass, David A.
16     REGISTRATION NUMBER: 38,153
17     REFERENCE/DOCKET NUMBER: 28967/33651
18     TELECOMMUNICATION INFORMATION:
19     TELEPHONE: 312/474-6300
20     TELEFAX: 312/474-0448
21     TELEX: 25-3856
22     INFORMATION FOR SEQ ID NO: 11:
23     SEQUENCE CHARACTERISTICS:
24     LENGTH: 415 amino acids
25     TYPE: amino acid
26     TOPOLOGY: linear
27     MOLECULE TYPE: protein
28     US-08-795-430-11

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:48:36 ; Search time 78.16 Seconds

(without alignments)
203.691 Million cell updates/sec

Title: US-09-427-657-2_COPY_211_419

Perfect score: 1218

Sequence: 1 CMSGKIDYRQVHSHIRSLP.....SYSEEVCRVPSYWKRPQMS 209

Scoring table:

BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1218	100.0	419	2	S69207
2	238	19.5	1700	2	S08167
3	180.5	14.8	160	2	J00542
4	151	12.4	1810	1	A32230
5	146.5	12.0	473	2	A56175
6	139.5	11.5	1203	2	A49175
7	139.5	11.5	1746	1	S19694
8	139.5	11.5	2195	2	T34264
9	139.5	11.5	2471	2	A49128
10	138	11.3	647	2	A43902
11	137.5	11.3	2437	2	S42612
12	137.5	11.3	2524	2	A35844
13	136	11.2	782	2	A51625
14	136	11.2	2907	2	A57278
15	135.5	11.1	1187	2	T18355
16	135.5	11.1	4006	2	T09070
17	135	11.1	2918	2	A54105
18	134.5	11.0	1620	2	T27283
19	134	11.0	1220	2	A56136
20	134	11.0	2352	2	T30201
21	133.5	10.9	2180	2	T29764
22	132.5	10.9	565	2	T16408
23	132	10.8	2019	1	J01322
24	131.5	10.8	2871	2	A55624
25	131.5	10.8	3002	2	A47221
26	130	10.7	2201	2	A32160
27	129.5	10.6	1372	2	T25933
28	129	10.6	626	2	T27319
29	128	10.5	1531	2	T42218

30	128	10.5	2703	1	A24420	notch protein - fr
31	127.5	10.5	2321	2	S78549	notch3 protein - h
32	127.5	10.5	3635	2	T10053	laminin alpha 5 ch
33	127	10.4	2871	2	A55567	fibrillin-1 - bovi
34	127	10.4	4135	2	T42629	tenascin-X - bovin
35	126.5	10.4	497	2	T27827	hypothetical prote
36	126	10.3	2555	2	A40043	hypothetical prote
37	125.5	10.3	3462	2	T46914	hypothetical prote
38	125.5	10.3	1111	2	T26972	hypothetical prote
39	125.5	10.3	2531	2	T31070	notch homolog - se
40	125.5	10.3	2531	2	A46019	Notch-1 protein -
41	125.5	10.3	3020	2	A43932	mucin 2 precursor,
42	124	10.2	570	2	A4836	fibropellin C prec
43	124	10.2	3712	2	S18253	laminin alpha-1 ch
44	123.5	10.1	220	2	S29195	antistatin - Hydra
45	123	10.1	1106	2	T13938	gene shuttle craft

ALIGNMENTS

RESULT 1
S69207
Vascular endothelial growth factor C precursor - human
N:Alternate names: FLT4 ligand DHM
C:Species: Homo sapiens (man)
C>Date: 27-Apr-1996 #sequence_revision 01-Nov-1996 #text_change 08-Oct-1999
C:Accession: S69207; S61795; S71443; S69208; G02659
R:Joukov, V.; Pajusola, K.; Kaipainen, A.; Chilov, D.; Laitinen, I.; Kukk, E.; Saksel
EMBO J. 15, 1751, 1996
A>Title: Corrigendum: A novel vascular endothelial growth factor, VEGF-C, is a ligand
A:Reference number: S69207; MUID:96203094
A:Accession: S69207
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-419 <JOU>
A:Cross-references: EMBL:X94216; NID:G1177488; PIDN:CAA63907.1; PTD:E221096; PTD:G118
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1995
A>Note: ONLY a part of the translation is shown
R:Joukov, V.; Pajusola, K.; Kaipainen, A.; Chilov, D.; Laitinen, I.; Kukk, E.; Saksel
EMBO J. 15, 290-298, 1996
A>Title: A novel vascular endothelial growth factor, VEGF-C, is a ligand for the FLT4
A:Reference number: S61795; MUID:96178224
A:Accession: S61795
A>Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 70-419 <JOU1>
A>Note: this sequence has been revised in reference S69207
A:Accession: S71443
A:Molecule type: protein
A:Residues: 'X', 104-120 <JOU2>
R:Lee, J.; Gray, A.; Yuan, J.; Luoh, S.M.; Avraham, H.; Wood, W.I.
submitted to the EMBL Data Library, December 1995
A:Description: Vascular endothelial growth factor related protein (VRP): A ligand and
A:Reference number: S69208
A:Accession: S69208
A:Molecule type: mRNA
A:Residues: 1-419 <LEP>
A:Cross-references: EMBL:U43142; NID:G1150988; PIDN:AAA85214.1; PTD:G1150989
R:Morris, J.C.
submitted to the EMBL Data Library, May 1996
A:Reference number: H01557
A:Accession: G02659
A>Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-419 <MOB>
A:Cross-references: EMBL:U58111; NID:G1373426; PTDN:AAH02909.1; PTD:G1373427
C:Genetics:
A:Gene: GDB:VEGFC; VRP
A:Cross-references: GDB:3890883; OMIM:601528
F:1-12/Domain: signal sequence #status predicted <SIG>
F:13-102/Domain: propeptide #status predicted <PRO>

F:103-419/Product: vascular endothelial growth factor C #status experimental <MAT>

Query Match 100.0%; Score 1218; DB 2; Length 419;

Best Local Similarity 100.0%; Pred. No. 1.2e-87;

Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CMSKLDVYQVHSHIRSLPATLPQCAANKTCPTMYMNNHICRCLAOEDFMFSSDAD 60
 |||||
 DB 211 CMSKLDVYQVHSHIRSLPATLPQCAANKTCPTMYMNNHICRCLAOEDFMFSSDAD 270
 QY 61 DSTDGFHDICGPKKELDEFTCCQCVCRAGLRPASGPHKELDRNSCCQVCKNKLFPSCGA 120
 |||||
 DB 271 DSTDGFHDICGPKKELDEFTCCQCVCRAGLRPASGPHKELDRNSCCQVCKNKLFPSCGA 330
 QY 121 NREFDENTCQCVCKRTCPNPNPLNPGKCAECETSPQKLLGKKFHHOTCSGYRRPCTN 180
 |||||
 DB 331 NREFDENTCQCVCKRTCPNPNPLNPGKCAECETSPQKLLGKKFHHOTCSGYRRPCTN 390
 QY 181 ROKACEPGFSYSEVGCRCVPSYWKRPOMS 209
 |||||
 DB 391 ROKACEPGFSYSEVGCRCVPSYWKRPOMS 419

RESULT 2

S08167 Balbiant ring 3 protein - midge (Chironomus tentans)

C:Species: Chironomus tentans

C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Jul-2000

C:Accession: S08167

R:Paulsson, G.; Lendahl, U.; Galli, J.; Ericsson, C.; Wieslander, L.

J. Mol. Biol. 211, 331-349, 1990

A:Title: The balbiant ring 3 gene in Chironomus tentans has a diverged repetitive struct

A:Reference number: S08167; MUID:90172404

A:Accession: S08167

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-1700 <PAU>

A:Cross-references: GB:X52263; NID:g7057; PIDN:CAA36506.1; PID:g7058

C:Genetics:

A:Gene: BR3

A:Map position: 4

C:Superfamily: unassigned Balbiant ring proteins

Query Match 19.5%; Score 238; DB 2; Length 1700;

Best Local Similarity 25.4%; Pred. No. 3.8e-11;

Matches 62; Conservative 26; Mismatches 74; Indels 82; Gaps 9;

QY 20 PATLPQCAANK-----TCPTNTMNNHICRCLAOEDFMFSSDAGDGS 62
 |||||
 DB 1204 PATLPQCAANK-----TCPTNTMNNHICRCLAOEDFMFSSDAGDGS 1259
 QY 63 TTGFFHDICGPKKELDEFTCCQCVCRAG-----LRPASGPH 97
 |||||
 DB 1260 -----C-KTWMNDDEMCCQCVCKPCPGGCKGVKMNANNTSCCECPADKAKPASCGDK 1311
 QY 98 KELDRNSCCQVCKNKLFPSCGANREFDENTCQCVCKRT--CPNPNPLN-----P 145
 |||||
 DB 1312 KSWNDSCSCQCKSKMPCGCPNPNPLNPKDECKSATGNCPLAGOTWNSQTCQSCPT 1371
 QY 146 GRC-----ACECTSPQKLLGKKFHHOTCSGY-----RRPCTNROKACEPGFSYS 192
 |||||
 DB 1372 GRCGTAGQVWCSAKCKVCYCAQKKCKSPKTDWNSCSCQCKPKMRRPTG---GCNAGRTPWD 1428
 QY 193 EEVC 196
 |||||
 DB 1429 DATC 1432

RESULT 3

J00542 185K secretory protein - midge (Chironomus tentans) (fragment)

N:Alternate names: balbiant ring 3 protein

C:Species: Chironomus tentans

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 17-Mar-2000

C:Accession: J00542

R:Digman, S.S.; Case, S.T.

Gene 88, 133-140, 1990

A:Title: Balbiant ring 3 in Chironomus tentans encodes a 185-kDa secretory protein wh

A:Reference number: J00542; MUID:90269600

A:Accession: J00542

A:Molecule type: mRNA

A:Residues: 1-160 <DIG>

A:Cross-references: GB:M24160

A:Experimental source: salivary gland

C:Superfamily: unassigned Balbiant ring proteins

Query Match 14.8%; Score 180.5; DB 2; Length 160;

Best Local Similarity 26.4%; Pred. No. 2.1e-07;

Matches 48; Conservative 24; Mismatches 77; Indels 33; Gaps 9;

QY 33 CPTNMMNNHICRCLAOEDFMFSSDAGDSTGDFHDICGPKKELDEFTCCQCVCRAGLRPA 92
 |||||
 DB 3 CKSPTQMTDSKLLC-----ECSTPATCGSKQTWC-----EACQCLCPG--DK 45
 QY 93 SCGPHKELDRNSCCQVCKNKLFPSCGANREFDENTCQCVCKRT-----CPNPNPLN 145
 |||||
 DB 46 NCGNKKFFDPCSECKCKNN--PST--SPQVADADCECKCPKDKQKPGCGDGGQKWND 101
 QY 146 GRCACECTSPQKLLGKKFHHOTCSGYRRCTNRORAKACEPGFSYSEVGC--VPSTYWK 204
 |||||
 DB 102 RVSCGCPVPRPDC--TNGQIYININTAC--CGGIDKPSCPKQOITNMKTCDCPCPNMK 157
 QY 205 RP 206
 |||||
 DB 158 EP 159

RESULT 4

A32230

tenascin precursor - chicken

N:Alternate names: cytotactin; hexabrachion

M:Comatins: tenascin 190K; tenascin 200K

C:Species: Gallus gallus (chicken)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: A32230; B32230; B33379; C33379; S01292; A30903

R:Jones, F.S.; Hoffman, S.; Cunningham, B.A.; Egelman, G.M.

Proc. Natl. Acad. Sci. U.S.A. 86, 1905-1909, 1989

A:Title: A detailed structural model of cytotactin: protein homologies, alternative R

A:Reference number: A32230; MUID:89184536

A:Accession: A32230

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1810 <JON>

A:Cross-references: GB:J04519; NID:g211717; PIDN:AAA48745.1; PID:g211718

A:Accession: B32230

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1044,1318-1810 <J02>

A:Cross-references: GB:J04519

R:Spring, U.; Beck, K.; Chiquet-Ehrismann, R.

Cell 59, 325-334, 1989

A:Title: Two contrary functions of tenascin: dissection of the active sites by recomb

A:Reference number: A33379; MUID:90030407

A:Accession: A33379

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-204,'G',206-221,'A',223-380,'D',382-386,'H',388-444,'HN',447-450,'V',45

A:Cross-references: GB:M23121; NID:g212746; PIDN:AAA49086.1; PID:g212749

A:Accession: B33379

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

A:Molecule type: mRNA

A:Residues: 1-204,'G',206-221,'A',223-380,'D',382-386,'H',388-444,'HN',447-450,'V',45

A:Accession: C33379

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tré
 A:Molecule type: mRNA
 A:Residues: 1-204, 'G', 206-221, 'A', 223-380, 'D', 382-386, 'H', 388-444, 'HN', 447-450, 'V', 452-5
 A:Cross-references: GB:M23121
 R:Pearson, C.A.; Pearson, D.; Shibahara, S.; Hofsteenge, J.; Chiquet-Ehrismann, R.
 EMBL J. 7, 2977-2982, 1988
 A:Title: Tenascin: cDNA cloning and induction by TGF-beta.
 A:Reference number: S01292; MUID:89030589
 A:Accession: S01292
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 27-181, 'R', 183-204, 'G', 206-221, 'A', 223-380, 'D', 382-386, 'H', 388-444, 'HN', 447-
 A:Cross-references: EMBL:X08030
 A:Note: part of this sequence was confirmed by protein sequencing
 C:Superfamily: tenascin; EGF homology; fibrinogen beta/gamma homology; fibronectin type
 C:Keywords: alternative splicing; calcium binding; cell adhesion; duplication; extracellular
 F:1-22/Domain: signal sequence #status predicted <SIG>
 F:23-33/Domain: propeptide #status predicted <PRO>
 F:34-1810/Product: tenascin 230k #status predicted <MAT>
 F:223-249/Domain: EGF homology <EGF1>
 F:316-342/Domain: EGF homology <EGF>
 F:359-673/Domain: fibronectin type III repeat homology <FN3A>
 F:681-765/Domain: fibronectin type III repeat homology <FN3B>
 F:773-857/Domain: fibronectin type III repeat homology <FN3C>
 F:865-949/Domain: fibronectin type III repeat homology <FN3D>
 F:957-1037/Domain: fibronectin type III repeat homology <FN3E>
 F:1046-1128/Domain: fibronectin type III repeat homology <FN3F>
 F:1137-1219/Domain: fibronectin type III repeat homology <FN3G>
 F:1228-1310/Domain: fibronectin type III repeat homology <FN3H>
 F:1318-1399/Domain: fibronectin type III repeat homology <FN3I>
 F:1407-1487/Domain: fibronectin type III repeat homology <FN3J>
 F:1495-1575/Domain: fibronectin type III repeat homology <FN3K>
 F:1590-1798/Domain: fibrinogen beta/gamma homology <FBG>
 F:1734-1747/Domain: calcium binding #status predicted <CAB>

Query Match 12.4%; Score 151; DB 1; Length 1810;
 Best Local Similarity 24.2%; Pred. No. 0.00022;
 Matches 61; Conservative 21; Mismatches 78; Indels 92; Gaps 18;

OY 24 POCQANKTPTNYMNNHIC---RCLAQEDF-----MESSDAGDS--TDG---F 66
 DB 185 PNC--SEBACPRNCL--NRGLCVRAKICEEGFTGEDCSQARPCSDCQGVDCVCF 241
 OY 67 HDIGCPNLEDETC-----QCVCRAGLRPAASC---PKHEIDRNC---Q 106
 DB 242 EGYTBP--DCGELCPHCGGCHGRGVGRCVCHBEGFTGEDCNEPLCPNCHRGRCVDNE 299
 OY 107 CVCKNKLFPSSQCG---ANREPDENTC---QCVCRR-----TCPRNQLN---P 145
 DB 300 CVCDDGYTGEDGELICNDPCDRGRCTNGTCFCBEGYTGEDCGELTPCPNNGNGRCEN 359
 OY 146 GKCAC-----ECTES--POKCLIKGKKFHHTQSCY-----RRP 177
 DB 360 GLCVCHEGFVGDDCSQKRCPCPCNNRGRCVDGR--CVCHEGYLGEDCGELRCPDCHNRGR 418
 OY 178 CTNRKACEPGP 189
 DB 419 CINGQCVCDGEP 430

RESULT 5
 A56175
 adhesive plaque protein Mgf2 precursor - Mediterranean mussel
 C:Species: Mytilus galloprovincialis (Mediterranean mussel)
 C:Date: 27-Apr-1995 #sequence revision 03-Oct-1995 #text-change 11-Jan-2000
 C:Accession: A56175
 R:Inoue, K.; Takeuchi, Y.; Miki, D.; Odo, S.
 J. Biol. Chem. 270, 6698-6701, 1995
 A:Title: Mussel adhesive plaque protein gene is a novel member of epidermal growth factor
 A:Reference number: A56175; MUID:95204464
 A:Accession: A56175
 A:Molecule type: mRNA

A:Residues: 1-473 <INO>
 A:Cross-references: GB:D43794; MUID:9602767; PIDN:BA07852.1; PID:dl008438; PID:960276
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 C:Keywords: duplication
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:387-419/Domain: EGF homology <EGF1>
 F:429-460/Domain: EGF homology <EGF>
 F:23-36,43,56,75,382,424,455,468,473/Modified site: 3',4'-dihydroxyphenylalanine (Tyr

Query Match 12.0%; Score 146.5; DB 2; Length 473;
 Best Local Similarity 21.5%; Pred. No. 0.00019;
 Matches 67; Conservative 26; Mismatches 83; Indels 135; Gaps 17;

OY 22 TLPCQANKTPTNYMNNHICRCL-----AOEDFMF---SSDAGDD 61
 DB 150 TGPCEV--HACKPNCNKKRCFPDGTGYCRCVDGSGPTCEMCKPNCPSNGTC 207
 OY 62 SFDGFHD-----ICGPNKLEDETC-----QCVCRAGLRPAASC-- 94
 DB 208 SADKFGDYSCERPGYFPECEERYVCAFPNCKNGICSSDGSGLTRCKGYSGLPTKV 267
 OY 95 -----GPHKELDR-----NSCQVCCKNKLFPSSQCGAN-----REPDENT----- 128
 DB 268 NVCKRPTPCNNSGRVKNKSSYCKICGKGYSGPTCEGNCKPNCQNRGCTPDNSDGRK 327
 OY 129 COCV-----CK---RTCPRNQPLNPGKC-----ACECT----- 153
 DB 338 CRCVGYGKPTCEDPNCPTKPCNKGCKNNGRIYTCCKAYGNRGRHCTDKAVKPNPC 387
 OY 154 -----ESPCKLLGKKRP-----HHOTGCGYR-----RPTNRKACEPGESY 191
 DB 388 VSKRCCKNKGCTIMGKAYRCCKAYGGRHCTKRSYKKNPCASRCKNKGCTDKNGSY 447
 OY 192 SEVCRCVPST 202
 DB 448 ---VCKCARGY 455

RESULT 6
 A49175
 Notch B protein - mouse (fragment)
 N:Alternate names: Notch homolog
 C:Species: Mus musculus (house mouse)
 C:Date: 21-Jan-1994 #sequence revision 05-Jan-1996 #text-change 20-Sep-1999
 C:Accession: A49175; PH1570; S32113
 R:Lardelli, M.; Lendahl, U.
 Exp. Cell Res. 204, 364-372, 1993
 A:Title: Notch A and Notch B--two mouse Notch homologues coexpressed in a wide variety
 A:Reference number: A49175; MUID:93178563
 A:Accession: A49175
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-1203 <LAR>
 A:Cross-references: EMBL:X68279; NID:9287989; PIDN:CAA48340.1; PID:9287990
 A:Experimental source: embryo
 A:Note: sequence extracted from NCBI backbone (NCBI:126158)
 C:Comment: This protein has many EGF repeats and 11n-12/Notch repeats.
 C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol
 F:143-174/Domain: EGF homology <EGX1>
 F:482-513/Domain: EGF homology <EGF1>
 F:560-591/Domain: EGF homology <EGF>
 F:674-705/Domain: EGF homology <EGX2>
 F:712-743/Domain: EGF homology <EGF3>
 F:836-867/Domain: EGF homology <EGX3>

Query Match 11.5%; Score 139.5; DB 2; Length 1203;
 Best Local Similarity 23.8%; Pred. No. 0.0013;
 Matches 62; Conservative 16; Mismatches 68; Indels 115; Gaps 15;
 OY 33 CPTNYMNNHICRCL-----AOEDFMFSSDAGDD-----STDG--FHDIC----- 70

```

Db      73 COTNPNNGOYICTCPGCGYTGADCTEDVDCAAMANSNPCEACKCVMTDPAFHECLKGYA 132
      71 GPNKELD-----EETCO-----CVCRAGLRPASGPHKELDRNSCO---CV- 108
      133 GPRCEMDINECHSDPCQNATCKLIGFTCLCMPEFK-----GVHELEVENEQSNPCVN 188
      109 ---CKKKLFPSCGAGNREDETCQCVCKRTCPRNQPLNPGKC-----ACECTE---- 154
      189 NGQCVKVRNFQCLCPPGFTGVCQ- IDIDCSSTPCLNACACIDHPNGVECOCATGFTG 247
      155 -----SPQKCLLKKKKFHHQTC-----SCYRRPCT 179
      248 ILCDENIDNCDDPPC-----HHGQCQDGDIDSYTCINPWTGAIICSDQIDECYSSPCL 300
      180 N-----RQKACEPGFS 190
      301 NDGRCIDLVNGQCNCQPGTS 321

```

```

RESULT 7
S19694
tenascin precursor - pig
N:Alternate names: contactin; hexabrachion
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S19694
R:Nishii, T.; Weinstein, J.; Gillespie, W.M.; Paulson, J.C.
Eur. J. Biochem. 202, 643-648, 1991
A:Title: Complete primary structure of porcine tenascin. Detection of tenascin transcript
A:Reference number: S19694; MUID:92104189
A:Accession: S19694
A:Molecule type: mRNA
A:Residues: 1-1746 <NTS>
A:Cross-references: EMBL:X61599; NID:q2124; PIDN:CAA43796.1; PID:92125
C:Superfamily: tenascin; EGF homology; fibrinogen beta/gamma homology; fibronectin type
C:Keywords: alternative splicing; calcium binding; cell adhesion; duplication; extracellular
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-1746/Product: tenascin #status predicted <MNT>
F:346-372/Domain: EGF homology <EGF>
F:377-403/Domain: EGF homology <EGF>
F:622-703/Domain: fibronectin type III repeat homology <FN3>
F:711-793/Domain: fibronectin type III repeat homology <FN3>
F:802-884/Domain: fibronectin type III repeat homology <FN3>
F:892-976/Domain: fibronectin type III repeat homology <FN3>
F:984-1064/Domain: fibronectin type III repeat homology <FN3>
F:1073-1155/Domain: fibronectin type III repeat homology <FN3>
F:1164-1246/Domain: fibronectin type III repeat homology <FN3>
F:1254-1323/Domain: fibronectin type III repeat homology <FN3>
F:1343-1423/Domain: fibronectin type III repeat homology <FN3>
F:1431-1511/Domain: fibronectin type III repeat homology <FN3>
F:1526-1734/Domain: fibrinogen beta/gamma homology <FBS>
F:38,166,184,337,788,1034,1079,1121,1354/Binding site: carbohydrate (Asn) (covalent) #st

```

```

Query Match      11.5%; Score 139.5; DB 1; Length 1746;
Best Local Similarity 26.2%; Pred. No. 0.0017;
Matches 53; Conservative 17; Mismatches 69; Indels 63; Gaps 13;

Oy 29 ANKTCPTVMNNHHC---RCLAOEDFMFSSDAGDSDG--FHDICGPKNELDETCOC 83
      248 SRETPVPCSEHGRCVGRCCQCGF---AGEDCNBPCLCHNCHRGRCVENE---C 299
      84 VCRAGLRPASG---PHKELDRNSC---QCVCNKKLFPSCGAGNREDETCOC----- 129
      300 VODEFTEDEGCELCPDPCDFRGRCINGTCYCDGFEDEGCG--RLACPHGRCRRGCE 357
      130 ---QVCV-----KRTCPRNQPLNPGKC---ACECTESFQKCLLKKKFFHHQTCSCY 174
      358 EGQCVCDGFGAGADCSERRCP-SDCHNRGRCLDGRCECDG-----FEGEDCGEL 406
      175 RRP-----CTNRQKACEPG 188

```

```

Db      407 RCPGCGSGHRCVNGOCVCDEG 428

```

```

RESULT 8
T34264
hypothetical protein F46C8.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T34264
R:Wilcox, L.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid F46C8.
A:Reference number: 221497
A:Accession: T34264
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2195 <WTL>
A:Cross-references: EMBL:U41624; PIDN:AAA83316.1; CESP:F46C8.4
C:Genetics:
A:Gene: CESP:F46C8.4
A:Introns: 19/3; 61/3; 80/1; 159/1; 182/1; 208/2; 246/1; 308/1; 367/1; 422/1; 466/1;
629/3; 1721/3; 1767/3; 1811/1; 1874/1; 1920/1; 2011/2; 2068/3; 2117/1; 2161/1

```

```

Query Match      11.5%; Score 139.5; DB 2; Length 2195;
Best Local Similarity 23.7%; Pred. No. 0.002;
Matches 52; Conservative 34; Mismatches 82; Indels 51; Gaps 13;

```

```

Oy 26 COANKTCPTVMNNHHCRCCLAOEDFMS--SDAGDDSDGFDHICGPKNELDETCOC 83
      1570 GULTYITCGNSOCIDICCKCPNNOGALINGRCSMGNMNCNIQ---CGTNQICIDQSCQ- 1626
      84 VCRAGL--RPASGPHKELDRNSCOV-----CKNKLFPSCGAGNREDETCOC--- 131
      1627 -CRPYTQPGSC-----LDRCNCIDVEBSQNLNR---QCGMNVCIQDQCCQSG 1675
      132 --VCKRTC--PRNQLNP-----GKCAECTES---PQKCLK-GKKFHHQTCVYR 176
      1676 YLVLETCISDRNCNVQPSVDAISGCMNCGNNOVCIQDQCLCRNGYVAGPCTGTGR 1735
      177 PCT-----NRQKACEPGFSFEVCRVPSYKRRQ 207
      1736 NCQHVHPDMGNCQRCQGNNOVCIQDQCCQCRNGYVAGPCTGTGR 1774

```

```

RESULT 9
A49128
cell-fate determining gene Notch2 protein - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 13-Aug-1999
C:Accession: A49128
R:Weinmaster, G.; Roberts, V.J.; Lemke, G.
Development 116, 931-941, 1992
A:Title: Notch2: a second mammalian Notch gene.
A:Reference number: A49128; MUID:93202015
A:Accession: A49128
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-2471 <WEI>
A:Experimental source: Schwann cell
A>Note: sequence extracted from NCBI backbone (NCBI:P127811)
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol
F:264-295/Domain: EGF homology <EGX1>
F:799-830/Domain: EGF homology <EGP1>
F:877-908/Domain: EGF homology <EGX2>
F:1029-1060/Domain: EGF homology <EGE>
F:1067-1098/Domain: EGF homology <EGX3>
F:1153-1184/Domain: EGF homology <EGF>
F:1191-1222/Domain: EGF homology <EGX4>
F:1876-1908/Domain: ankyrin repeat homology <AN1>
F:1909-1941/Domain: ankyrin repeat homology <AN2>
F:1943-1975/Domain: ankyrin repeat homology <AN3>
F:1976-2008/Domain: ankyrin repeat homology <AN4>

```


F:2009-2041/Domain: ankyrin repeat homology <AN5>

Query Match 11.5%; Score 139.5; DB 2; Length 2471;

Best Local Similarity 23.8%; Pred. No. 0.0022; Matches 62; Conservative 16; Mismatches 68; Indels 115; Gaps 15;

```

QY 33 CPTNMMNNHICRL-----AOEDFMSSDAGD-----STDG-FHDIC----- 70
DB 390 CPTNPLNGOYICTPQATKGAADCTEDVDECAAMNSNPCHEACKCVNTDGAFCCECLKGYA 449
QY 71 GNPKELD-----EETCO-----CYCRAGLRPASCGPKHELDNRSCQ---CY- 108
DB 450 GPRCEMDINECHSDPCQDNATCLDKIGFTCLCMGPK-----GVHCELEVNQCQSNPCVN 505
QY 109 ---CKNKLFPSSCGANREPDENTCQCVCCKRTCPRNQPLNPGKC-----ACECTE---- 154
DB 506 NGQCVDKVNRFCQCLCPPTGTPVCO-IDIDDCSSTPCNLGAKCIDHPNGYECQCATGFTG 564
QY 155 -----SPQCKLKGKKFHHQTC-----SCYRRPCT 179
DB 565 TLCDENIDNCDPDP-----HHGQCQDGDIDSYTCINCPYMGALCSDQIDECTSSPCL 617
QY 180 N-----RQKACEPGFS 190
DB 618 NDGRCIDLVNGYQCNQCPCTS 638

```

RESULT 10

A43902
 A:Species: eastern newt (fragment)
 C:Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 12-Feb-1999
 C:Accession: A43902
 R:Ronda, H.; Poulin, M.L.; Taasava, R.A.; Chiu, I.M.
 Dev. Biol. 148, 219-232, 1991
 A:Title: Characterization of a new tenascin cDNA and localization of tenascin mRNA during
 A:Reference number: A43902; MUID:92038434
 A:Accession: A43902
 A:Molecule type: mRNA
 A:Residues: 1-647 <OND>
 A:Cross-references: GB:M76615
 A>Note: sequence extracted from NCBI backbone (NCBIN:64543, NCBI:64547)
 C:Superfamily: tenascin; EGF homology; fibronogen beta/gamma homology; fibronectin type
 C:Keywords: extracellular matrix; glycoprotein; tandem repeat
 F:287-313/Domain: EGF homology <EGF>
 F:346-427/Domain: fibronectin type III repeat homology <FN1>
 F:435-517/Domain: fibronectin type III repeat homology <FN2>

Query Match 11.3%; Score 138; DB 2; Length 647;

Best Local Similarity 28.2%; Pred. No. 0.0011; Matches 55; Conservative 11; Mismatches 73; Indels 56; Gaps 13;

```

QY 33 CPTNMMNNHICRLAODEFMFSSDAGDSDTDGFHDICGNKELDERTC---QCVCRAGL 89
DB 39 CPTNMMNNHICRLAODEFMFSSDAGDSDTDGFHDICGNKELDERTC---QCVCRAGL 89
QY 39 CPNNC---NNRGRVCYDEYQDEGFTGDCSE---LIC-PNCCFPRGRINCINVCCTDEGF 91
DB 39 CPNNC---NNRGRVCYDEYQDEGFTGDCSE---LIC-PNCCFPRGRINCINVCCTDEGF 91
QY 90 RPASCG---PKKELDRNSC---QCVCKNKLFPSSCG---ANREPDENTC---QCVCKR 135
DB 90 RPASCG---PKKELDRNSC---QCVCKNKLFPSSCG---ANREPDENTC---QCVCKR 135
QY 92 TGEDGELTCCPNMNNRGRVCYGLVCDDGDFGQDSCSELRCPNCDNRGRVCYNGKVCCKE 151
DB 92 TGEDGELTCCPNMNNRGRVCYGLVCDDGDFGQDSCSELRCPNCDNRGRVCYNGKVCCKE 151
QY 136 -----TCPRNQPLNPGKA---CECTESPOKCLKGKKFHHQTCSCYRRP--- 177
DB 136 -----TCPRNQPLNPGKA---CECTESPOKCLKGKKFHHQTCSCYRRP--- 177
QY 178 ---CTNRQKACEPGF 189
DB 201 RGRVCYNGQCVQDEGF 215

```

RESULT 11

S42612

transmembrane protein precursor - zebra fish

C:Species: Brachydanio rerio (zebra fish)
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Sep-1999

C:Accession: S42612
 R:Bierkamp, C.; Campos-Ortega, J.A.
 Mech. Dev. 43, 87-100, 1993

A:Title: A zebrafish homologue of the Drosophila neurogenic gene Notch and its patter
 A:Reference number: S42612; MUID:94128602

```

A:Accession: S42612
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2437 <BIE>
A:Cross-references: EMBL:X69088; NID:q433866; PIDN:CA448831.1; PID:q433867
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol
F:755-786/Domain: EGF homology <EGF1>
F:1023-1054/Domain: EGF homology <EGF2>
F:1185-1216/Domain: EGF homology <EGF3>
F:1915-1947/Domain: ankyrin repeat homology <AN1>
F:1948-1980/Domain: ankyrin repeat homology <AN2>
F:1982-2014/Domain: ankyrin repeat homology <AN3>
F:2015-2047/Domain: ankyrin repeat homology <AN4>
F:2048-2080/Domain: ankyrin repeat homology <AN5>

```

Query Match 11.3%; Score 137.5; DB 2; Length 2437;
 Best Local Similarity 22.9%; Pred. No. 0.0031;
 Matches 57; Conservative 13; Mismatches 66; Indels 113; Gaps 15;

```

QY 61 DSTDGFHDICG-----NKEIDE-----ETC-----QCVCRAGLRPASC-GP 96
DB 655 DKINGEYCEYCEPYSGSMCNINIDCALNPNCHNGGICIDGVNSFCLCPDPRDACLISO 714
QY 97 HKELDRNSC-----QCVCKNKLFPSSCGAN-REPDENTC-----Q 130
DB 715 HNECSSNPCHSGCLDQINSYRCVCEAGMGRNCDININECLSNPCVNGGCTCKDMTSGYL 774
QY 131 CVCK-----RTPRNQPLNPGKC-----ACECTE-----SPQ 157
DB 775 CTRAGFSGNPCQNMNINEASNPCLNGSCIDVAGFKNCMLPTGCEVENVLAPCSR 834
QY 158 KCLLKG-----KKFHHQTC-----SCYRRPCTNRQKACE---PGFSYSE 193
DB 835 PCKNGYCVRESDEPFSFCNCPAGWQGTCEVDINECVANPCTN-GGYCENLRGFGQ--- 890
QY 194 EYCRVPSY 202
DB 891 ---CRCPGR 897

```

RESULT 12

A35844

Xotch protein - African clawed frog

C:Species: Xenopus laevis (African clawed frog)
 C:Date: 12-Oct-1990 #sequence_revision 12-Oct-1990 #text_change 13-Aug-1999

C:Accession: A35844
 R:Coffman, C.; Harris, W.; Kintner, C.
 Science 249, 1438-1441, 1990

A:Title: Xotch, the Xenopus homolog of Drosophila notch.
 A:Reference number: A35844; MUID:90385285

```

A:Accession: A35844
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 1-2524 <COF>
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol
C:Keywords: transmembrane protein
F:146-177/Domain: EGF homology <EGX1>
F:184-215/Domain: EGF homology <EGF1>
F:222-254/Domain: EGF homology <EGF2>
F:456-487/Domain: EGF homology <EGX2>
F:757-788/Domain: EGF homology <EGF3>
F:1025-1056/Domain: EGF homology <EGX3>
F:1924-1956/Domain: ankyrin repeat homology <AN1>
F:1957-1989/Domain: ankyrin repeat homology <AN2>

```



```
Db 295 ICKNLEHSEILYRLKNGOVNLLLETELELHVADSPVQEGKEPQCGSLKETEECDCE 354
Oy 84 VCRAGLRPASCGPHRELDNSQCYCKNKLFPSCGANREPDENTCQCV-----CKR--- 135
Db 355 ACKC-----OECEN---CSCSE---LTGC---OEAFCSCAQEHGCGQEEESC 393
Oy 136 TCPRNQPLNPGKCA-----CECTESPOKCLKGGKFHHQTCSCYRRPCTNRQACEPGFS 190
Db 394 ACPNT-----TCACTEHECECTES-----TCGCENMPCECEEFACD----- 429
Oy 191 YSEEVCRCV 199
Db 430 CSEEHCECV 438
```

Search completed: October 17, 2001, 14:48:38
Job time: 284 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:51:40 ; Search time 42.92 Seconds

(without alignments)
166.808 Million cell updates/sec

Title: US-09-427-657-2_COPY_211_419

Perfect score: 1218

Sequence: 1 CMSKLDVYRQVHSIIRRLP.....SYSEWCVCPYSWKRPQMS 209

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1218	100.0	419	1	VEGC_HUMAN
2	1086	89.2	415	1	VEGC_MOUSE
3	238	19.5	1700	1	BAR3_CHITE
4	148	12.2	1808	1	TENA_CHICK
5	146.5	12.0	473	1	PP2_MYTGA
6	140.5	11.5	2482	1	VWF_PIG
7	140.5	11.5	2813	1	VWF_CANPA
8	139.5	11.5	1746	1	TENA_PIG
9	137.5	11.3	2437	1	NOTC_BRARE
10	137.5	11.3	2524	1	NOTC_XENLA
11	136	11.2	2907	1	FBM2_MOUSE
12	135	11.1	2911	1	FBM2_HUMAN
13	131.5	10.8	2871	1	FBM1_HUMAN
14	131.5	10.8	2871	1	FBM1_MOUSE
15	130	10.7	931	1	EMR1_MOUSE
16	130	10.7	2201	1	TENA_HUMAN
17	128.5	10.6	1696	1	PCK5_BRACL
18	128	10.5	2703	1	NOTC_DROME
19	127.5	10.5	2444	1	NTC1_HUMAN
20	127.5	10.5	3635	1	LMA5_MOUSE
21	127	10.4	2871	1	FBM1_BOVIN
22	125.5	10.3	2531	1	NTC1_MOUSE
23	125.5	10.3	5179	1	MG22_HUMAN
24	125	10.3	769	1	ITB2_PIG
25	124	10.2	570	1	FBP3_STRPU
26	124	10.2	3712	1	LMA_DROME
27	123.5	10.1	220	1	ANTA_HYDMA
28	123	10.1	1106	1	STC_DROME
29	123	10.1	2318	1	NTC3_MOUSE
30	123	10.1	2813	1	VWF_HUMAN
31	123.5	10.1	1877	1	PCK5_MOUSE
32	121	9.9	1680	1	FUR2_DROME
33	120.5	9.9	379	1	WIFI_HUMAN

34	120.5	9.9	2139	1	CRB_DROME	P10040 drosophila
35	120	9.9	787	1	ITB3_MOUSE	O54890 mus musculus
36	119	9.8	2531	1	NTC1_RAT	O07008 rattus norv
37	118.5	9.7	1964	1	NTC4_MOUSE	P31695 mus musculus
38	118	9.7	1104	1	NFX1_HUMAN	Q12986 homo sapien
39	118	9.7	3672	1	LML2_CAEEL	Q21313 caenorhabd
40	118	9.7	4289	1	TENX_HUMAN	P22105 homo sapien
41	117.5	9.6	379	1	WIFI_MOUSE	O9wua1 mus musculus
42	117.5	9.6	1168	1	LMB3_MOUSE	O61087 mus musculus
43	116.5	9.6	1403	1	NID2_MOUSE	O88322 mus musculus
44	115.5	9.5	1394	1	TGFB_HUMAN	P22064 homo sapien
45	115.5	9.5	4544	1	LRL1_HUMAN	O07954 homo sapien

ALIGNMENTS

RESULT 1

ID	VEGC_HUMAN	STANDARD	PRT	419 AA.
AC	P49767:			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR (VEGF-C) (VASCULAR			
DE	ENDOTHELIAL GROWTH FACTOR RELATED PROTEIN) (VRP) (FLT4 LIGAND) (FLT4-			
DE	L)			
GN	VEGFC.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RM	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 103-120.			
RX	MEDLINE=96178224; PubMed=8617204;			
RA	Joukov V., Pajusola K., Kaipainen A., Chllov D., Lahtinen I., Kukk E.,			
RA	Saksela O., Kalkkinen N., Alitalo K.;			
RT	"A novel vascular endothelial growth factor: VEGF-C, is a ligand for			
RT	the Flt4 (VEGFR-3) and KDR (VEGFR-2) receptor tyrosine kinases.";			
RL	EMBO J. 15:290-298(1996).			
RM	[2]			
RP	ERRATUM.			
RX	MEDLINE=96203094; PubMed=8612600;			
RA	Joukov V., Pajusola K., Kaipainen A., Chllov D., Lahtinen I., Kukk E.,			
RA	Saksela O., Kalkkinen N., Alitalo K.;			
RL	EMBO J. 15:1751-1751(1996).			
RM	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96312526; PubMed=8700872;			
RA	Lee J., Gray A., Yuan J., Luoh S.-M., Avraham H., Wood W. I.;			
RT	"Vascular endothelial growth factor-related protein: a ligand and			
RT	specific activator of the tyrosine kinase receptor Flt4.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 93:1988-1992(1996).			
RM	[4]			
RP	SEQUENCE FROM N.A.			
RA	Fitz L., Morris J.C., Towler P.S., Long A.J., Greco R.,			
RA	Burgess P., Gianotti J., Charleita A., Hennessey D., Kovacic S.,			
RA	Fitzgerald M., Scaltreito H., Welch N., Neben S., Finerly H.,			
RA	Zollner R., Wang J., Nickbarg E., Gassaway R., Turner K.,			
RL	Wood C.R.;			
CC	Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.			
CC	- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL			
CC	CELL GROWTH.			
CC	- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.			
CC	- PTM: PROBABLY PROTEOLITICALLY PROCESSED IN THE C-TERMINUS.			
CC	- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb.ch/announce/			

or send an email to license@isb-sib.ch.

```
CC DR EMBL: X94216; CAA63907.1; -
CC DR EMBL: U43142; AA85214.1; -
CC DR EMBL: U58111; AA802909.1; -
CC DR HSSP: P15692; 1VPF.
CC DR MIM: 601528; -.
CC DR InterPro: IPR000072; -.
CC DR InterPro: IPR002400; -.
CC DR Pfam: PF00341; PDGF_1;
CC DR PROSITE: PS00249; PDGF_1;
CC DR PROSITE: PS50278; PDGF_2; 1.
CC DR Mitogen: Growth factor; Glycoprotein; Signal; Repeat.
CC FT SIGNAL 1 102 POTENTIAL.
CC FT PROPEP 2 419 POTENTIAL.
CC FT CHAIN 103 419 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
CC FT DOMAIN 275 365 4 X 24 AA TANDEM REPEATS.
CC FT REPEAT 275 298 1.
CC FT REPEAT 299 322 2.
CC FT REPEAT 323 346 3.
CC FT REPEAT 347 365 4 (PARTIAL).
CC FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SO SEQUENCE 419 AA; 46883 MW; 9F598719DB3E014F CRC64;
```

Query Match 100.0%; Score 1218; DB 1; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.3e-94;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 CMSKLDVYRQVHSIIIRSLPATLPQQAANKTCPTNYMNNHICRLAODEMFSSDAGD 60
DB 211 CMSKLDVYRQVHSIIIRSLPATLPQQAANKTCPTNYMNNHICRLAODEMFSSDAD 270
OY 61 DSTDGFHDICGNPKELDEFTCCQVCRAGLRPASCGPHKELDRNSCCQVCKNKLFPSSQGA 120
DB 271 DSTDGFHDICGNPKELDEFTCCQVCRAGLRPASCGPHKELDRNSCCQVCKNKLFPSSQGA 330
OY 121 NEEFDENTCCQVCKRTCPRPNOPLNPGKACCECTESPQKLLGKKFHHQTCSCYRRPCTN 180
DB 331 NEEFDENTCCQVCKRTCPRPNOPLNPGKACCECTESPQKLLGKKFHHQTCSCYRRPCTN 390
OY 181 ROKACEPGFSSEYECRCVPSTYWKRPQMS 209
DB 391 ROKACEPGFSSEYECRCVPSTYWKRPQMS 419
```

RESULT 2
VEGC_MOUSE STANDARD; PRT; 415 AA.

```
ID VEGC_MOUSE STANDARD; PRT; 415 AA.
AC P97953;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR (VEGF-C) (FLT4 LIGAND)
DE (FLT4-L).
GN VEGFC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAUB/C;
RX MEDLINE=97164697; PubMed=9012504;
RA Kukk E., Lymboussaki A., Taira S., Kaipainen A., Jeltsch M.,
RA Jonkov V., Altaio K.;
RT "VEGF-C receptor binding and pattern of expression with VEGF-R3
RL suggests a role in lymphatic vascular development.";
RN [2]
```

RP SEQUENCE FROM N.A.
RC STRAIN=BAUB/C;
RX MEDLINE=97388482; PubMed=9247316;
RA Filz L.J., Morris J.C., Towler P., Long A., Burgess P., Greco R.,
RA Wang J., Gassaway R., Nickbarg E., Kovacic S., Charletta A.,
RA Gianocci J., Finerty H., Zollner R., Beler D.R., Leak L.V.,
RA Turner K.J., Wood C.R.;
RT "Characterization of murine Flt4 ligand/VEGF-C";
RL Oncogene 15:613-618(1997).
CC -!- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
CC CELL GROWTH.
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -!- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

```
CC DR EMBL: U73620; AAC52984.1; -
CC DR EMBL: U58112; AAB46707.1; -
CC DR HSSP: P15692; 1VPF.
CC DR MGD: MGI:109124; Vegfc.
CC DR InterPro: IPR000072; -.
CC DR InterPro: IPR002400; -.
CC DR Pfam: PF00341; PDGF_1;
CC DR PRINTS: PR00458; GFCYSKNOT.
CC DR PROSITE: PS00249; PDGF_1; 1.
CC DR PROSITE: PS50278; PDGF_2; 1.
CC DR Mitogen: Growth factor; Glycoprotein; Signal; Repeat.
CC FT SIGNAL 1 98 POTENTIAL.
CC FT PROPEP 2 419 POTENTIAL.
CC FT CHAIN 99 98 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
CC FT DOMAIN 271 361 4 X 24 AA TANDEM REPEATS.
CC FT REPEAT 271 294 1.
CC FT REPEAT 295 318 2.
CC FT REPEAT 319 342 3.
CC FT REPEAT 343 361 4 (PARTIAL).
CC FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SO SEQUENCE 415 AA; 46471 MW; D9D3DD3CECC659D6 CRC64;
```

Query Match 89.2%; Score 1086; DB 1; Length 415;
Best Local Similarity 85.6%; Pred. No. 1.2e-83;
Matches 179; Conservative 18; Mismatches 12; Indels 0; Gaps 0;

```
OY 1 CMSKLDVYRQVHSIIIRSLPATLPQQAANKTCPTNYMNNHICRLAODEMFSSDAGD 60
DB 207 CMSKLDVYRQVHSIIIRSLPATLPQQAANKTCPTNYMNNHICRLAODEMFSSDAD 266
OY 61 DSTDGFHDICGNPKELDEFTCCQVCRAGLRPASCGPHKELDRNSCCQVCKNKLFPSSQGA 120
DB 267 DSTDGFHDICGNPKELDEFTCCQVCRAGLRPASCGPHKELDRNSCCQVCKNKLFPSSQGA 326
OY 121 NEEFDENTCCQVCKRTCPRPNOPLNPGKACCECTESPQKLLGKKFHHQTCSCYRRPCTN 180
DB 327 NEEFDENTCCQVCKRTCPRPNOPLNPGKACCECTENTQKFLGKKFHHQTCSCYRRPCAN 386
OY 181 ROKACEPGFSSEYECRCVPSTYWKRPQMS 209
DB 387 ROKACEPGFSSEYECRCVPSTYWKRPQMS 415
```

RESULT 3
BAR3_CHITE STANDARD; PRT; 1700 AA.
ID BAR3_CHITE
AC Q03376;
DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE BALBIANI RING PROTEIN 3 PRECURSOR.
 GN BR3.
 OS Chironomus tentans (Midge).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
 OC Chironomidae; Chironomidae; Chironominae; Chironomus.
 NX NCBI_TaxID=7153;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=90172404; PubMed=1689777;
 RA Paulsson G., Lendahl U., Galli J., Ericsson C., Wieslander L.;
 RT "The Balbiani ring 3 gene in Chironomus tentans has a diverged
 RT repetitive structure split by many introns."
 RL J. Mol. Biol. 211:331-349(1990).
 CC -1- FUNCTION: USED BY THE LARVAE TO CONSTRUCT A SUPRAMOLECULAR
 CC STRUCTURE, THE LARVAL TUBE. BALBIANI RING PROTEIN 3 COULD PLAY A
 CC ROLE AS A TRANSPORT PROTEIN THAT BINDS TO OTHER PROTEINS.
 CC INTRACELLULARLY AND IN THE GLAND LUMEN IN ORDER TO PREVENT THESE
 CC FROM FORMING WATER-INSOLUBLE FIBERS TOO EARLY.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- TISSUE SPECIFICITY: SALIVARY GLAND.
 CC -1- DOMAIN: HAS 82 APPROXIMATE REPEATS OF CYS-X-CYS-X-CYS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: X52263; CAA36506.1;
 DR PIR: S08167; S08167.
 DR HSSP: P18055; 2MRB.
 DR InterPro: IPR000853;
 DR PRINTS: PR00876; MTNEMATODE.
 DR Repeat: Signal.
 KW SIGNAL
 FT CHAIN 1 1700 POTENTIAL.
 FT SIGNAL 1 1700 BALBIANI RING PROTEIN 3.
 SQ SEQUENCE 1700 AA; 186145 MW; 34202B28521B0815 CRC64;

Query Match 19.5%; Score 238; DB 1; Length 1700;
 Best Local Similarity 25.4%; Pred. No. 1,2e-12;
 Matches 62; Conservative 26; Mismatches 74; Indels 82; Gaps 9;

QY 20 PATLPQCCQANK-----TCPTNYMNNHICRCLAQEDFMFSSPAGDGS 62
 DB 1204 PPAFPCSNNOQYSNVSCGCGNPGKPKNGCPGNOIWCNCRCKPKMKERPADN---- 1259
 QY 63 TGGFHDICPKNELDEETQCVCYCRAG-----LRPASCGRH 97
 DB 1260 -----C-KTWMMNDQMCQVCCKPCPCGCGKGVKMNANNTSCCECPADKAKPASCGRK 1311
 QY 98 KELDNRSCQCVCKNKLFPSCGAGNEFDENTCQVCYCKPT--CPRNOPLN-----P 145
 DB 1312 KSMNDSCSCQCKSKMPCGCGCPNNOQMKDECKCSATGNCPCAGOTWNSQTCQCSQAPT 1371
 QY 146 GRC-----ACRCESPOKCLLGGKKFHQTCSCY---RPCTNRQAKACEPGFSYS 192
 DB 1372 GKCTGAQVWCSCAKCKVCPAQKCKDSPKTDWNSCSCQCPKMMRPPTG---GCNMGRTWD 1428
 QY 193 EYVC 196
 DB 1429 DATC 1432

RESULT 4
 TENA_CHICK
 ID TENA_CHICK STANDARD: PRT: 1808 AA.

AC P10039; P13132; 073584; 073585;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE TENASCIN PRECURSOR (TN) (HEXABRACHION) (CYTOTACTIN) (NEURONECTIN)
 DE (GMEB) (JI) (MIOENDINCOUS ANTIGEN) (GLIOMA-ASSOCIATED-EXTRACELLULAR
 DE MATRIX ANTIGEN) (GP 150-225).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=90030407; PubMed=2478295;
 RA Spring J., Beck K., Chiquet-Ehrismann R.;
 RT "Two contrary functions of tenascin: dissection of the active sites
 RT by recombinant tenascin fragments."
 RL Cell 59:325-334(1989).
 RN [2]
 RP SEQUENCE OF 27-722 FROM N.A., AND SEQUENCE OF 79-96.
 RC TISSUE=Fibroblast;
 RX MEDLINE=89030589; PubMed=2460335;
 RA Pearson C.A., Pearson D., Shibahara S., Hofsteenge J.,
 RA Chiquet-Ehrismann R.;
 RL "Tenascin: cDNA cloning and induction by TGF-beta."
 RL EMBL J. 7:2977-2982(1988).
 RN [3]
 RP SEQUENCE OF 464-1018 AND 1412-1661 FROM N.A., AND SEQUENCE OF 852-868.
 RC TISSUE=Embryo;
 RX MEDLINE=88176910; PubMed=2451243;
 RA Jones F.S., Burgoon M.P., Hoffman S., Crossin K.L., Cunningham B.A.,
 RA Edelman G.M.;
 RT "A cDNA clone for cytotactin contains sequences similar to epidermal
 RT growth factor-like repeats and segments of fibronectin and
 RT fibrinogen."
 RL Proc. Natl. Acad. Sci. U.S.A. 85:2186-2190(1988).
 CC -1- FUNCTION: SAW (SUBSTRATE-ADHESION MOLECULE) THAT APPEARS TO
 CC INHIBIT CELL MIGRATION. MAY PLAY A ROLE IN SUPPORTING THE GROWTH
 CC OF EPITHELIAL TUMORS.
 CC -1- SUBUNIT: HEXAMERIC. A HOMOTRIMER MAY BE FORMED IN THE TRIPLE
 CC COILED-COIL REGION AND MAY BE STABILIZED BY DISULFIDE RINGS AT
 CC BOTH ENDS. TWO OF SUCH HALF-HEXABRACHIONS MAY BE DISULFIDE LINKED
 CC WITHIN THE CENTRAL GLOBULE.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: 230 KDA (SHOWN HERE), 200 KDA
 CC AND 190 KDA; ARE PRODUCED BY ALTERNATIVE SPLICING. THEY ARE
 CC PRODUCED IN A TISSUE-AND TIME-SPECIFIC MANNER DURING DEVELOPMENT.
 CC -1- INDUCTION: BY TGF-BETA.
 CC -1- SIMILARITY: CONTAINS 13.5 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 11 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS A DOMAIN RELATED TO THE C-TERMINAL PART OF
 CC THE BETA AND GAMMA CHAINS OF FIBRINOGEN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: M23121; AAA49086.1;
 DR EMBL: X08031; CAA40811.1;
 DR EMBL: X08030; CAA30824.1; ALT_TERM.
 DR EMBL: J03641; AAA48748.1; ALT_SEQ.
 DR EMBL: M20816; AAA48749.1; ALT_SEQ.
 DR PIR: A30903; A30903.
 DR PIR: A31930; A31930.
 DR PIR: A33379; A33379.
 DR PIR: B33379; B33379.
 DR PIR: C33379; C33379.

Query Match	Best Local Similarity	Score 148:	DB 1:	Length 1808:
Matches 61:	Conservative 20:	Mismatches 79:	Indels 92:	Gaps 18:
QY 24 PQCAANTCTPTNNMNNHTC-----RCLAQDEP-----MFSSDAGDGS--TDG---F 66	185 PNC--SEPAKRNCL--NRGLCLVGRGKCLCEBGFEBDCSQAACPSDCNDQKCVGAVCYCF 241	12.28:		
QY 67 HDICSPNKELDEPCTC-----QCVCRAGLRPASCG--PKKELDRNSC--Q 106	242 EGYTGP--DCEBELCPHCCGTHGRGVGRVCYCHGEGFTGEDNEPLRCPRNCHNRGRVDNE 299			
QY 107 CVCKNKLTPPSQCG---ANREPDENTC---QCVCKR-----TCPRNPOLN---P 145	300 CVCDEGYTGEDCGELICPNDCFDRGRCLNTGTCFCEGYTGEDCGELICPNNCNGNRCEN 359			
QY 146 GKAC-----ECTES--POKCLLKGGKFFHQTGSCY-----RRP 177	360 GLCVCHGEFVGDDCSQKRCPRKDCNNRCHGVADGR--CVCHGEGLGEDGELRCRPNDCNNRGR 418			
QY 178 CTNRKACEPGF 189	419 CINGCCVCDSEGF 430			
DB 419 CINGCCVCDSEGF 430				
RESULT 5				
FP2_MYTGA ID	FP2_MYTGA	STANDARD:	PRT:	473 AA.
AC Q25464;				
DT 30-MAY-2000 (Rel. 39, Created)				
DT 30-MAY-2000 (Rel. 39, Last sequence update)				
DT 01-OCT-2000 (Rel. 40, Last annotation update)				
DE ADHESIVE PLAQUE MATRIX PROTEIN 2 PRECURSOR (FOOT PROTEIN 2) (MGFP2)				
GN FP2.				
NC Mytilus galloprovincialis (Mediterranean mussel).				
OC Eukaryota: Metazoa: Mollusca: Bivalvia: Pteriomorpha: Mytiloida:				

FT	DISULFID	270	281	BY SIMILARITY.
FT	DISULFID	275	290	BY SIMILARITY.
FT	DISULFID	292	301	BY SIMILARITY.
FT	DISULFID	306	317	BY SIMILARITY.
FT	DISULFID	311	328	BY SIMILARITY.
FT	DISULFID	330	339	BY SIMILARITY.
FT	DISULFID	346	357	BY SIMILARITY.
FT	DISULFID	351	366	BY SIMILARITY.
FT	DISULFID	368	377	BY SIMILARITY.
FT	DISULFID	387	399	BY SIMILARITY.
FT	DISULFID	393	408	BY SIMILARITY.
FT	DISULFID	410	419	BY SIMILARITY.
FT	DISULFID	429	440	BY SIMILARITY.
FT	DISULFID	434	449	BY SIMILARITY.
FT	DISULFID	451	460	BY SIMILARITY.
SQ	SEQUENCE	473 AA;	51772 MW; BAVBAGCBA49A0F CRC64;	
 Query Match 12.0%; Score 146.5; DB 1; Length 473; Best Local Similarity 21.5%; Pred. No. 1.8e-05; Matches 67; Conservative 26; Mismatches 83; Indels 135; Gaps				
OY	22 TLPCQAKNTCTPTNYMNM-----NHICRCL-----AOEDFMF---SSDAGDD	61		
Dd	150 TGPRCEV--HACKRPCKNNKGRCEPDKTKYKKRCVDGISGPLCDENACKPPNCNGTTC	207		
OY	62 STDFHD-----ICGPKNELDEETC-----OCYCRAGLRASC--	94		
Dd	208 SADRFDYSCERGFPGPCERYVCAPNPKNGJCSDSGGYCRCKGGYSPTKV	267		
OY	95 -----GHKEIDR-----NSCCVCANKLFPSSCGAN-----REPENT-----	128		
Dd	268 NVCKPTCKNSGRNVNKSSSYNICIGGVSGPGCNVCKPNPCQNRCRGYPDNDSDGEK	327		
OY	129 CQCY-----CK----RTCPRNQLPMPKC-----ACECT-----	153		
Dd	328 CRCVGGYKRGPCCEKPPPCMTKFCCKNGCKCNYGKITTYTCKCAWGRGRCTDKATKPNC	387		
OY	154 -----ESPCKLCGKRKF-----HHOTCSGYR-----RCTNRORACEPGSFY	191		
Dd	388 VSKPRCKNRKGCIWNGKAYRCMCAGYGGRHCTKRSYKKNPCASRPCRKNGCKTDKNGY	447		
OY	192 SEEYCRGVPSY	202		
Dd	448 ---VCKCARGY	455		
 RESULT 6 VWF_PIG STANDARD; PRT; 2482 AA. ID VWF_PIG O28833; DT 01-OCT-2000 (Rel. 40, Created) DT 01-OCT-2000 (Rel. 40, Last sequence update) DT 01-OCT-2000 (Rel. 40, Last annotation update) DE VON WILLEBRAND FACTOR PRECURSOR (VWF) (FRAGMENT). GN FBVMF OR VWF. OS Sus scrofa (Pig). OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus. CX NCBI_TaxId=9823; RN [1] RP SEQUENCE FROM N.A. RA Seaman W.T., Read M.S., Bellinger D.A., Nichols T.C.; RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases. RN [2] RP SEQUENCE OF 397-553 FROM N.A. RX MEDLINE=93356762; PubMed=8352759; RA Laveigne J.M., Plao Y.C., Ferreira V., Kerblirou-Nabias D., RA Bahnak B.R., Meyer D.; RT "Primary structure of the factor VIII binding domain of human, porcine and rabbit von Willebrand factor"; BT Biochem. Biophys. Res. Commun. 194:1019-1024(1993). CC -!- FUNCTION: IMPORTANT IN THE MAINTENANCE OF HOMEOSTASIS, IT				

PARTICIPATES IN PLATELET-VESSEL WALL INTERACTIONS BY FORMING A
NONCOVALENT COMPLEX WITH COAGULATION FACTOR VIII AT THE SITE OF
VASCULAR INJURY (BY SIMILARITY).

-1- SUBUNIT: MULTIMERIC (BY SIMILARITY).

-1- PIV: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR
INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).

-1- SIMILARITY: CONTAINS 3 WMFC DOMAINS.

-1- SIMILARITY: CONTAINS A C-TERMINAL CYSTEINE KNOT-LIKE DOMAIN (CTCK).

-1- SIMILARITY: SOME: TO SILKMOH HEMOCYTIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; AF052036; AAC06229.1; -
CC EMBL; S64541; AAB27829.2; -
CC HSSP; P04275; IATZ.
CC InterPro; IPR000359; -
CC InterPro; IPR001007; -
CC InterPro; IPR001846; -
CC InterPro; IPR002035; -
CC InterPro; IPR002919; -
CC Pfam; PF00007; Cys_knot; 1.
CC Pfam; PF00092; wva; 3.
CC Pfam; PF00093; wvc; 3.
CC Pfam; PF00094; wvd; 3.
CC Pfam; PF01826; TIL; 3.
CC PRINTS; PR00365; ENDOTHELIN.
CC PRINTS; PR00453; VMFADOMAIN.
CC PROSITE; PS00234; VMFA; 3.
CC PROSITE; PS01208; VMFC; 3.
CC PROSITE; PS01185; CTCK_1; 1.
CC PROSITE; PS01225; CTCK_2; 1.
CC Blood coagulation; Platelet; Glycoprotein; Extracellular matrix;
KW Plasma; Endothelial cell; Repeat; Cell adhesion.
FT NON_TER 1
FT CHAIN <1 437 BY SIMILARITY.
FT DOMAIN 438 2482 VON WILLEBRAND FACTOR.
FT DOMAIN 62 215 AMINO-TERMINAL.
FT DOMAIN 438 461 VMFA 2.
FT DOMAIN 462 507 E1.
FT DOMAIN 500 527 CX.
FT DOMAIN 541 687 VMFA 3.
FT DOMAIN 947 1127 VMFA 1.
FT DOMAIN 1167 1334 VMFA 2.
FT DOMAIN 1360 1540 VMFA 3.
FT DOMAIN 1619 1771 VMFA 4.
FT DOMAIN 1885 1930 E2.
FT DOMAIN 1924 1997 VMFC 1.
FT DOMAIN 2098 2164 VMFC 2.
FT DOMAIN 2249 2319 VMFC 3.
FT DOMAIN 2393 2481 CTCK.
FT SITE 2176 2178 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 441 482 BY SIMILARITY.
FT DISULFID 450 478 BY SIMILARITY.
FT DISULFID 484 495 BY SIMILARITY.
FT DISULFID 541 670 BY SIMILARITY.
FT DISULFID 563 705 BY SIMILARITY.
FT DISULFID 572 667 BY SIMILARITY.
FT DISULFID 588 758 BY SIMILARITY.
FT DISULFID 734 758 BY SIMILARITY.
FT DISULFID 745 785 BY SIMILARITY.
FT DISULFID 763 765 BY SIMILARITY.
FT DISULFID 827 839 BY SIMILARITY.
FT DISULFID 833 843 BY SIMILARITY.
FT DISULFID 800 804 BY SIMILARITY.
FT DISULFID 870 873 BY SIMILARITY.
FT DISULFID 908 911 BY SIMILARITY.
FT DISULFID 942 1128 BY SIMILARITY.

FT DISULFID 1338 1339 BY SIMILARITY.
FT DISULFID 1355 1341 BY SIMILARITY.
FT DISULFID 1548 1573 BY SIMILARITY.
FT DISULFID 1568 1609 OR 1611 (BY SIMILARITY).
FT DISULFID 1641 1792 BY SIMILARITY.
FT DISULFID 1619 1754 BY SIMILARITY.
FT DISULFID 1596 1757 BY SIMILARITY.
FT DISULFID 1662 1670 BY SIMILARITY.
FT DISULFID 2393 2443 BY SIMILARITY.
FT DISULFID 2408 2457 BY SIMILARITY.
FT DISULFID 2419 2473 BY SIMILARITY.
FT DISULFID 2423 2475 BY SIMILARITY.
FT DISULFID ? 2480 BY SIMILARITY.
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 905 905 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1184 1184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1243 1243 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1640 1640 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1892 1892 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1959 1959 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2026 2026 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2069 2069 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2215 2215 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2254 2254 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2459 2459 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 2482 AA; 272394 MW; D499B7DDFBBCAEDD CRC64;

Query Match 11.5%; Score 140.5; DB 1; Length 2482;
Best Local Similarity 24.7%; Pred. No. 0.00022;
Matches 64; Conservative 21; Mismatches 93; Indels 81; Gaps 16;

QY 1 CMSKIDVYROYHSIRSRSLPATLPQC-----QAANKTCPTVMNNHICRC-LAQE 50
DB 1952 CLSC-----RRVNCITLQPPCTPAPACGICEVARKQEAHQCCP-----ETECVDLVSC 2001
QY 51 DF--WFSSDAGDSDTGFHDICGPNKEDEEFTCCQ--VCRAGLRPASCGPHKE----- 99
DB 2002 DLPPVPHCEGGLOPLTNPBGECRPV-----FTCACRKECECRGPLP-SCPHRRPALUKT 2055
QY 100 --LDNNSQCVCYCKNKLFPSCGCANREFDENICQCYCKRTC-----PRNDPLN 144
DB 2056 QCCDEYECACNCVNTLSCPLGYLASTVINDCGCT-TTTCLEPDKVCVHRTGYVPGQFWE 2114
QY 145 PGKACECTESPQCKLTKGKFFHHQTCGYRPPCNRKACEPGPSY---SEEV-RCVP 200
DB 2115 EGCVCYCTCTD-----LDAVWGLRVACCAQKPC---EDSCRPQFTYVLAHSGCCGKCLP 2165
QY 201 -----SYWK 204
DB 2166 SACKVIGSFRDSDSVSYWK 2184

RESULT 7
ID VWF_CANFA STANDARD; PRT; 2813 AA.
AC Q28295; Q28311; Q9TS14;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE VON WILLEBRAND FACTOR PRECURSOR (VWF).
GN F8VWF OR VWF.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
CX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA stoy S.J., Shibuya H., Nonneman D.J., Holzhauser J., Mohammed I.H.,
RJ Johnson G.S.,
RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

RA Montgomery R.R., Fahs S., Montgomery M.W.;
RN Submitted (Aug-1996) to the EMBL/Genbank/DBJ databases.
[3]
RE SEQUENCE FROM N.A.
RA Venta P.J., Li J., Yuzbasliyan-Gurkan V., Brewer G.J., Schall W.D.;
RT "Complete sequence of the structural gene for canine von Willebrand
factor and identification of a mutation causing Scottish terrier von
Willebrand's disease."
RL Submitted (Oct-1998) to the EMBL/Genbank/DBJ databases.
[4]
RP SEQUENCE OF 1234-1669 FROM N.A.
RC TISSUE=Blood;
RA Mancuso D.J., Christopher P.A., Kloner P.A., Montgomery R.R.;
RT "The canine von Willebrand factor gene: sequence and expression of
a region encoding the glycoprotein Ib/IX binding domain."
RL Submitted (Jan-1994) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: IMPORTANT IN THE MAINTENANCE OF HOMEOSTASIS. IT
PARTICIPATES IN PLATELET-VESSEL WALL INTERACTIONS BY FORMING A
NONCOVALENT COMPLEX WITH COAGULATION FACTOR VIII AT THE SITE OF
VASCULAR INJURY (BY SIMILARITY).
CC -1- SUBUNIT: MULTIMERIC. (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: BLOOD.
CC -1- PTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR
INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 3 VMFA DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 VMFC DOMAINS.
CC -1- SIMILARITY: CONTAINS 4 VMFD DOMAINS.
CC -1- SIMILARITY: CONTAINS A C-TERMINAL CYSTEINE KNOT-LIKE DOMAIN (CTCK).
CC -1- SIMILARITY: SOME, TO SILKWOOM HEMOCYTIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L76227; AAB0549.1; -;
DR EMBL: AF099154; AAD04919.1; -;
DR EMBL: U66246; AAB93766.1; -;
DR HSSP: P04275; IAT2.
DR InterPro: IPR000359; -;
DR InterPro: IPR001007; -;
DR InterPro: IPR001846; -;
DR InterPro: IPR002035; -;
DR InterPro: IPR002919; -;
DR Pfam: PF000007; Cys_Knot; 1.
DR Pfam: PF01826; TIL; 4.
DR Pfam: PF00092; vwa; 3.
DR Pfam: PF00093; vwc; 3.
DR Pfam: PF00094; vwd; 4.
DR PRINTS: PR00453; VWFADOMAIN.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS0234; VMFA; 3.
DR PROSITE: PS01208; VMFC; 3.
DR PROSITE: PS01208; VMFC; 3.
KW Blood coagulation; Platelet; Glycoprotein; Extracellular matrix;
KW Signal; Endothelial cell; Repeat; Cell adhesion; Signal.
FT SIGNAL 1 22
FT PROPEP 23 763
FT CHAIN 764 2813
FT DOMAIN 35 179
FT DOMAIN 388 541
FT DOMAIN 764 787
FT DOMAIN 788 833
FT DOMAIN 826 853
FT DOMAIN 867 1013
FT DOMAIN 1277 1453
FT DOMAIN 1498 1665
FT DOMAIN 1691 1871
FT DOMAIN 1950 2102
VMFD 1.
VMFD 2.
AMINO-TERMINAL.
EL.
CX.
VMFD 3.
VWFA 1.
VWFA 2.
VWFA 3.
VWFA 4.

FT DOMAIN 2216 2261 E2.
FT DOMAIN 2255 2326 VMFC 1.
FT DOMAIN 2429 2495 VMFC 2.
FT DOMAIN 2580 2650 VMFC 3.
FT DOMAIN 2724 2812 CTCK.
FT SITE 531 533 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 698 700 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 2507 2509 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 767 808 BY SIMILARITY.
FT DISULFID 776 804 BY SIMILARITY.
FT DISULFID 810 821 BY SIMILARITY.
FT DISULFID 867 996 BY SIMILARITY.
FT DISULFID 889 1031 BY SIMILARITY.
FT DISULFID 898 993 BY SIMILARITY.
FT DISULFID 914 921 BY SIMILARITY.
FT DISULFID 1060 1084 BY SIMILARITY.
FT DISULFID 1071 1111 BY SIMILARITY.
FT DISULFID 1089 1091 BY SIMILARITY.
FT DISULFID 1153 1165 BY SIMILARITY.
FT DISULFID 1149 1169 BY SIMILARITY.
FT DISULFID 1126 1130 BY SIMILARITY.
FT DISULFID 1196 1199 BY SIMILARITY.
FT DISULFID 1234 1237 BY SIMILARITY.
FT DISULFID 1272 1458 BY SIMILARITY.
FT DISULFID 1669 1670 BY SIMILARITY.
FT DISULFID 1686 1872 BY SIMILARITY.
FT DISULFID 1879 1904 BY SIMILARITY.
FT DISULFID 1899 1940 OR 1942 (BY SIMILARITY).
FT DISULFID 1972 2123 BY SIMILARITY.
FT DISULFID 1950 2085 BY SIMILARITY.
FT DISULFID 1927 2088 BY SIMILARITY.
FT DISULFID 1993 2001 BY SIMILARITY.
FT DISULFID 2724 2774 BY SIMILARITY.
FT DISULFID 2739 2768 BY SIMILARITY.
FT DISULFID 2750 2804 BY SIMILARITY.
FT DISULFID 2754 2806 BY SIMILARITY.
FT DISULFID ? 2811 BY SIMILARITY.
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 211 211 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 666 666 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 857 857 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1231 1231 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1515 1515 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1574 1574 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2223 2223 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2290 2290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2357 2357 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2400 2400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2546 2546 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2585 2585 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2790 2790 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 55 55 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 70 70 V -> E (IN REF. 2).
FT CONFLICT 266 266 A -> I (IN REF. 3).
FT CONFLICT 280 280 I -> V (IN REF. 2).
FT CONFLICT 409 411 VCH -> ICG (IN REF. 2).
FT CONFLICT 994 994 G -> A (IN REF. 1).
FT CONFLICT 1021 1021 F -> L (IN REF. 2).
FT CONFLICT 2381 2381 L -> P (IN REF. 2).
FT CONFLICT 2406 2406 P -> L (IN REF. 2).
SQ SEQUENCE 2813 AA; 309716 MM; 5DP93E1E5E72F60C CRC64;

Query Match 11.5%; Score 140.5; DB 1; Length 2813;
Best Local Similarity 20.5%; Pred. No. 0.00024;
Matches 67; Conservative 37; Mismatches 56; Indels 167; Gaps 21;

QY 32 TCPTNYMNN--HICRCLAQEDPMFSSDAGDSDSTGFIHICGPNKELDETC-----Q 82
DB 2202 SCPPSYVNHCHGCRCEGN--TSSCGDQPSBEC--FCPPNOVMLEGSCVPEACTQ 2256
QY 83 CVCRAGLR-----PA-----SCGPHK--ELDRNS 104


```

FT DISULFID 377 387 BY SIMILARITY.
FT DISULFID 381 392 BY SIMILARITY.
FT DISULFID 394 403 BY SIMILARITY.
FT DISULFID 408 418 BY SIMILARITY.
FT DISULFID 422 423 BY SIMILARITY.
FT DISULFID 425 434 BY SIMILARITY.
FT DISULFID 439 449 BY SIMILARITY.
FT DISULFID 443 454 BY SIMILARITY.
FT DISULFID 456 465 BY SIMILARITY.
FT DISULFID 470 480 BY SIMILARITY.
FT DISULFID 474 485 BY SIMILARITY.
FT DISULFID 487 496 BY SIMILARITY.
FT DISULFID 501 511 BY SIMILARITY.
FT DISULFID 505 516 BY SIMILARITY.
FT DISULFID 518 527 BY SIMILARITY.
FT DISULFID 532 542 BY SIMILARITY.
FT DISULFID 536 547 BY SIMILARITY.
FT DISULFID 549 558 BY SIMILARITY.
FT DISULFID 563 573 BY SIMILARITY.
FT DISULFID 567 578 BY SIMILARITY.
FT DISULFID 580 589 BY SIMILARITY.
FT DISULFID 594 604 BY SIMILARITY.
FT DISULFID 598 609 BY SIMILARITY.
FT DISULFID 611 620 BY SIMILARITY.
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 788 788 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1034 1034 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1079 1079 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1121 1121 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1354 1354 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPPLIC 1072 1162 MISSING (IN MAJOR ISOFORM).
FT VARSPPLIC 1072 1253 MISSING (IN MINOR-1 ISOFORM).
FT CONFLICT 1007 1007 T->M (IN REF. 2).
SQ SEQUENCE 1746 AA; 191399 MW; 56549B1CFE5E5C88 CRC64;

```

Query Match 11.5%; Score 139.5; DB 1; Length 1746;
 Best Local Similarity 26.2%; Pred. No. 0.0002;
 Matches 53; Conservative 17; Mismatches 69; Indels 63; Gaps 13;

```

QY 29 ANKTCPTNMMNNHC---RCLAEDEPFMSDAGDSTDG--FHDICGPNKEDEPTCC 83
DB 248 SKETGVPCSEHGRCVGRVCQDEGF-----AGEDCNPEPLHNGHRCVENB---C 299
QY 84 VCRAGLRPASCG---PKKELDRNSC---QCVCKNKLFPSCGANEFEDENTC----- 129
DB 300 VDEGFTGDEGCELCPCFKCFCRGRCINCTGCDSEGFEGEDCG--RLACPHCRGRGRCE 357
QY 130 --QCVC-----KRCPRNOLNPGKC--ACECTESPOKCLLKGRKHNOTGSCY 174
DB 358 ESOQCVCDDEGFAGADCSERKCP--SDCHNNGRCILDGRCECDG-----FEGEDCGEL 406
QY 175 RRP-----CTNRKACEPG 188
DB 407 KCPGGCSGHRGVNGQCVCDG 428

```

```

RESULT 9
ID NOTC_BRARE STANDARD: PRT: 2437 AA.
AC P46530:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN PRECURSOR.
GN NOTCH.
OS Brexyciano rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Rasbora; Danio.

```

```

OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=embryo;
RA MEDLINE=94128602; PubMed=8297791;
RT Bierkamp C., Campos-Ortega J.A.;
RT "A zebrafish homologue of the Drosophila neurogenic gene Notch and
RL Mech. Dev. 43:87-100(1993).
CC - FUNCTION: IMPLICATED IN CELL FATE SPECIFICATIONS DURING
CC EMBRYO DEVELOPMENT. MAY BE INVOLVED IN THE FORMATION OF THE
CC NEURAL PLATE, NOTOCHORD AND BRAIN VESICLES.
CC - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC - DEVELOPMENTAL STAGE: EXPRESSED IN ALL CELLS IN PREGASTRULATION
CC STAGES. DURING GASTRULATION IS DIFFERENTIALLY EXPRESSED,
CC ACCUMULATING PREDOMINANTLY IN THE PRECHORDAL MESODERM AND
CC NOTOCHORD. AT THE END OF GASTRULATION, EXPRESSED ALONG THE
CC ANTERIOR-POSTERIOR AXIS INCLUDING THE DEVELOPING NEURAL PLATE
CC AND DIFFERENTIATING MESODERM. ALSO PRESENT IN THE DEVELOPING
CC BRAIN AND HEAD REGIONS.
CC - SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
CC - SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC - SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC - SIMILARITY: CONTAINS 6 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X69088; CAA48831.1; -.
DR HSSP: P00740; IIXA.
DR ZFIN: ZDB-GENE-990415-173; notch.
DR InterPro: IPR000152; -.
DR InterPro: IPR0000561; -.
DR InterPro: IPR0000800; -.
DR InterPro: IPR001336; -.
DR InterPro: IPR001438; -.
DR InterPro: IPR001881; -.
DR InterPro: IPR002110; -.
DR Pfam: PF000008; EGF_36.
DR Pfam: PF000023; ank; 6.
DR Pfam: PF000066; notch; 3.
DR PRINTS: PR00009; EGFTGF.
DR PRINTS: PR00010; EGFBLOD.
DR PROSITE: PS50088; ANK_REPEAT; 4.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; 23.
DR PROSITE: PS00022; EGF_1; 34.
DR PROSITE: PS01187; EGF_CA; 22.
DR Transmembrane: Signal; Glycoprotein.
KV Transmembrane: Signal; Glycoprotein.
FT SIGNAL 1 20
FT CHAIN 21 2437
FT DOMAIN 21 1724
FT TRANSMEM 1725 1747
FT DOMAIN 1748 2437
FT DOMAIN 21 57
FT DOMAIN 58 98
FT DOMAIN 101 138
FT DOMAIN 139 175
FT DOMAIN 177 215
FT DOMAIN 217 254
FT DOMAIN 256 292
FT DOMAIN 294 332
FT DOMAIN 334 370
FT DOMAIN 371 409
FT DOMAIN 411 449
FT DOMAIN 451 487

```



```

FT DISULFID 536 551 BY SIMILARITY.
FT DISULFID 553 562 BY SIMILARITY.
FT DISULFID 569 579 BY SIMILARITY.
FT DISULFID 574 588 BY SIMILARITY.
FT DISULFID 590 599 BY SIMILARITY.
FT DISULFID 606 617 BY SIMILARITY.
FT DISULFID 611 626 BY SIMILARITY.
FT DISULFID 628 637 BY SIMILARITY.
FT DISULFID 644 654 BY SIMILARITY.
FT DISULFID 649 663 BY SIMILARITY.
FT DISULFID 665 674 BY SIMILARITY.
FT DISULFID 681 692 BY SIMILARITY.
FT DISULFID 703 712 BY SIMILARITY.
FT DISULFID 719 729 BY SIMILARITY.
FT DISULFID 724 738 BY SIMILARITY.
FT DISULFID 740 749 BY SIMILARITY.
FT DISULFID 756 767 BY SIMILARITY.
FT DISULFID 761 776 BY SIMILARITY.
FT DISULFID 778 787 BY SIMILARITY.
FT DISULFID 794 805 BY SIMILARITY.
FT DISULFID 799 814 BY SIMILARITY.
FT DISULFID 816 825 BY SIMILARITY.
FT DISULFID 832 843 BY SIMILARITY.
FT DISULFID 837 854 BY SIMILARITY.
FT DISULFID 856 865 BY SIMILARITY.
FT DISULFID 872 883 BY SIMILARITY.
FT DISULFID 877 892 BY SIMILARITY.
FT DISULFID 894 903 BY SIMILARITY.
FT DISULFID 910 921 BY SIMILARITY.
FT DISULFID 915 930 BY SIMILARITY.
FT DISULFID 932 941 BY SIMILARITY.
FT DISULFID 986 997 BY SIMILARITY.
FT DISULFID 991 1006 BY SIMILARITY.
FT DISULFID 1008 1017 BY SIMILARITY.
FT DISULFID 1024 1035 BY SIMILARITY.
FT DISULFID 1029 1044 BY SIMILARITY.
FT DISULFID 1046 1055 BY SIMILARITY.
FT DISULFID 1062 1073 BY SIMILARITY.
FT DISULFID 1067 1082 BY SIMILARITY.
FT DISULFID 1084 1093 BY SIMILARITY.
FT DISULFID 1100 1121 BY SIMILARITY.
FT DISULFID 1115 1130 BY SIMILARITY.
FT DISULFID 1132 1141 BY SIMILARITY.
FT DISULFID 1148 1159 BY SIMILARITY.
FT DISULFID 1153 1168 BY SIMILARITY.
FT DISULFID 1170 1179 BY SIMILARITY.
FT DISULFID 1186 1197 BY SIMILARITY.
FT DISULFID 1191 1206 BY SIMILARITY.
FT DISULFID 1208 1217 BY SIMILARITY.
FT DISULFID 1224 1243 BY SIMILARITY.
FT DISULFID 1237 1252 BY SIMILARITY.

Query Match 11.3% Score 137.5; DB 1; Length 2524..
Best Local Similarity 21.7% Pred. No. 0.00039;
Matches 55; Conservative 21; Mismatches 77; Indels 101; Gaps 12;

```

```

Oy 191 --YSEVCRCPVSY 202
Db 999 DGINFTCCCPGPF 1012

RESULT 11
FBN2_MOUSE STANDARD; PRT; 2907 AA.
ID 061555; 063957;
AC 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
GN FIBRILLIN 2 PRECURSOR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RP MEDLINE=95263670; PubMed=7744963;
RA Zhang H., Hu W., Ramirez F.;
RT "Developmental expression of fibrillin genes suggests heterogeneity
of extracellular microfibrils."
RL J. Cell Biol. 129:1165-1176(1995).
[2]
RP SEQUENCE OF 210-317 FROM N.A.
RX MEDLINE=94140368; PubMed=8307578;
RA Li X., Pereira L., Zhang H., Sanginetti C., Ramirez F., Bonadio J.,
RA Francke U.;
RT "Fibrillin genes map to regions of conserved mouse/human synteny on
mouse chromosomes 2 and 18."
RL Genomics 18:667-672(1993).
CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC -!- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 7 TGF-BETA BINDING-PROTEIN DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
-----
DR EMBL; L39790; AAA74908.1; -
DR EMBL; S69359; AAC60685.1; -
DR MGI; MGI:95490; Fbn2.
DR InterPro: IPR000152; -
DR InterPro: IPR000561; -
DR InterPro: IPR000822; -
DR InterPro: IPR001438; -
DR InterPro: IPR001881; -
DR InterPro: IPR002212; -
DR Pfam; PF00008; EGF; 46.
DR Pfam; PF00683; TB; 9.
DR PRINTS; PR00010; EGFBLD.
DR PROSITE; PS00010; ASX_HYDROXYL; 43.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 36.
DR PROSITE; PS01187; EGF_CA; 43.
KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW Repeat; Signal; Multigene family.
FT SIGNAL 1 28
FT CHAIN 29 2907
FT DOMAIN 111 142 EGF-LIKE 1, NON-CALCIUM BINDING.
FT DOMAIN 145 176 EGF-LIKE 2, NON-CALCIUM BINDING.
FT DOMAIN 176 208 EGF-LIKE 3, NON-CALCIUM BINDING.
FT DOMAIN 276 317 EGF-LIKE 4, CALCIUM-BINDING.
FT DOMAIN 318 359 EGF-LIKE 5, CALCIUM-BINDING.

```


FT	REPEAT	360	426	TGFBP 1.
FT	DOMAIN	487	527	EGF-LIKE 6, NON-CALCIUM BINDING.
FT	DOMAIN	528	567	EGF-LIKE 7, CALCIUM-BINDING.
FT	DOMAIN	568	609	EGF-LIKE 8, CALCIUM-BINDING.
FT	DOMAIN	610	650	EGF-LIKE 9, CALCIUM-BINDING.
FT	DOMAIN	651	691	EGF-LIKE 10, CALCIUM-BINDING.
FT	REPEAT	692	760	TGFBP 2.
FT	DOMAIN	761	802	EGF-LIKE 11, CALCIUM-BINDING.
FT	DOMAIN	803	844	EGF-LIKE 12, CALCIUM-BINDING.
FT	DOMAIN	845	883	EGF-LIKE 13, CALCIUM-BINDING.
FT	DOMAIN	884	989	EGF-LIKE 14, CALCIUM-BINDING.
FT	REPEAT	990	1065	TGFBP 3.
FT	DOMAIN	1066	1107	EGF-LIKE 15, CALCIUM-BINDING.
FT	DOMAIN	1108	1150	EGF-LIKE 16, CALCIUM-BINDING.
FT	DOMAIN	1151	1192	EGF-LIKE 17, CALCIUM-BINDING.
FT	DOMAIN	1193	1234	EGF-LIKE 18, CALCIUM-BINDING.
FT	DOMAIN	1235	1275	EGF-LIKE 19, CALCIUM-BINDING.
FT	DOMAIN	1276	1317	EGF-LIKE 20, CALCIUM-BINDING.
FT	DOMAIN	1318	1359	EGF-LIKE 21, CALCIUM-BINDING.
FT	DOMAIN	1360	1400	EGF-LIKE 22, CALCIUM-BINDING.
FT	DOMAIN	1401	1441	EGF-LIKE 23, CALCIUM-BINDING.
FT	DOMAIN	1442	1483	EGF-LIKE 24, CALCIUM-BINDING.
FT	DOMAIN	1484	1524	EGF-LIKE 25, CALCIUM-BINDING.
FT	DOMAIN	1525	1565	EGF-LIKE 26, CALCIUM-BINDING.
FT	REPEAT	1566	1642	TGFBP 4.
FT	DOMAIN	1643	1684	EGF-LIKE 27, CALCIUM-BINDING.
FT	DOMAIN	1685	1726	EGF-LIKE 28, CALCIUM-BINDING.
FT	REPEAT	1727	1800	TGFBP 5.
FT	DOMAIN	1801	1842	EGF-LIKE 29, CALCIUM-BINDING.
FT	DOMAIN	1843	1884	EGF-LIKE 30, CALCIUM-BINDING.
FT	DOMAIN	1885	1926	EGF-LIKE 31, CALCIUM-BINDING.
FT	DOMAIN	1927	1965	EGF-LIKE 32, CALCIUM-BINDING.
FT	DOMAIN	1966	2008	EGF-LIKE 33, CALCIUM-BINDING.
FT	DOMAIN	2009	2048	EGF-LIKE 34, CALCIUM-BINDING.
FT	DOMAIN	2049	2090	EGF-LIKE 35, CALCIUM-BINDING.
FT	REPEAT	2091	2163	TGFBP 6.
FT	DOMAIN	2164	2205	EGF-LIKE 36, CALCIUM-BINDING.
FT	DOMAIN	2206	2245	EGF-LIKE 37, CALCIUM-BINDING.
FT	DOMAIN	2246	2286	EGF-LIKE 38, CALCIUM-BINDING.
FT	DOMAIN	2287	2330	EGF-LIKE 39, CALCIUM-BINDING.
FT	DOMAIN	2331	2372	EGF-LIKE 40, CALCIUM-BINDING.
FT	REPEAT	2373	2441	TGFBP 7.
FT	DOMAIN	2442	2483	EGF-LIKE 41, CALCIUM-BINDING.
FT	DOMAIN	2484	2524	EGF-LIKE 42, CALCIUM-BINDING.
FT	DOMAIN	2525	2563	EGF-LIKE 43, CALCIUM-BINDING.
FT	DOMAIN	2564	2606	EGF-LIKE 44, CALCIUM-BINDING.
FT	DOMAIN	2607	2646	EGF-LIKE 45, CALCIUM-BINDING.
FT	DOMAIN	2647	2687	EGF-LIKE 46, CALCIUM-BINDING.
FT	DOMAIN	2688	2727	EGF-LIKE 47, CALCIUM-BINDING.
FT	DISULFID	115	130	BY SIMILARITY.
FT	DISULFID	119	134	BY SIMILARITY.
FT	DISULFID	132	141	BY SIMILARITY.
FT	DISULFID	149	159	BY SIMILARITY.
FT	DISULFID	153	164	BY SIMILARITY.
FT	DISULFID	166	175	BY SIMILARITY.
FT	DISULFID	180	190	BY SIMILARITY.
FT	DISULFID	184	196	BY SIMILARITY.
FT	DISULFID	198	207	BY SIMILARITY.
FT	DISULFID	280	292	BY SIMILARITY.
FT	DISULFID	287	301	BY SIMILARITY.
FT	DISULFID	303	316	BY SIMILARITY.
FT	DISULFID	322	334	BY SIMILARITY.
FT	DISULFID	329	343	BY SIMILARITY.
FT	DISULFID	345	358	BY SIMILARITY.
FT	DISULFID	491	503	BY SIMILARITY.
FT	DISULFID	498	512	BY SIMILARITY.
FT	DISULFID	514	526	BY SIMILARITY.
FT	DISULFID	532	542	BY SIMILARITY.
FT	DISULFID	537	551	BY SIMILARITY.
FT	DISULFID	553	566	BY SIMILARITY.
FT	DISULFID	572	584	BY SIMILARITY.
FT	DISULFID	579	593	BY SIMILARITY.
FT	DISULFID	595	608	BY SIMILARITY.
FT	DISULFID	614	625	BY SIMILARITY.
FT	DISULFID	620	634	BY SIMILARITY.
FT	DISULFID	636	649	BY SIMILARITY.
FT	DISULFID	655	666	BY SIMILARITY.
FT	DISULFID	661	675	BY SIMILARITY.
FT	DISULFID	677	690	BY SIMILARITY.
FT	DISULFID	755	777	BY SIMILARITY.
FT	DISULFID	772	786	BY SIMILARITY.
FT	DISULFID	788	801	BY SIMILARITY.
FT	DISULFID	807	819	BY SIMILARITY.
FT	DISULFID	814	828	BY SIMILARITY.
FT	DISULFID	830	843	BY SIMILARITY.
FT	DISULFID	849	859	BY SIMILARITY.
FT	DISULFID	854	868	BY SIMILARITY.
FT	DISULFID	870	883	BY SIMILARITY.
FT	DISULFID	952	964	BY SIMILARITY.
FT	DISULFID	959	973	BY SIMILARITY.
FT	DISULFID	975	988	BY SIMILARITY.
FT	DISULFID	1070	1082	BY SIMILARITY.
FT	DISULFID	1077	1091	BY SIMILARITY.
FT	DISULFID	1093	1106	BY SIMILARITY.
FT	DISULFID	1112	1124	BY SIMILARITY.
FT	DISULFID	1119	1133	BY SIMILARITY.
FT	DISULFID	1135	1149	BY SIMILARITY.
FT	DISULFID	1155	1167	BY SIMILARITY.
FT	DISULFID	1162	1176	BY SIMILARITY.
FT	DISULFID	1178	1191	BY SIMILARITY.
FT	DISULFID	1197	1209	BY SIMILARITY.
FT	DISULFID	1204	1218	BY SIMILARITY.
FT	DISULFID	1220	1233	BY SIMILARITY.
FT	DISULFID	1239	1250	BY SIMILARITY.
FT	DISULFID	1246	1259	BY SIMILARITY.
FT	DISULFID	1261	1274	BY SIMILARITY.
FT	DISULFID	1280	1292	BY SIMILARITY.
FT	DISULFID	1287	1301	BY SIMILARITY.
FT	DISULFID	1302	1316	BY SIMILARITY.
FT	DISULFID	1322	1334	BY SIMILARITY.
FT	DISULFID	1329	1343	BY SIMILARITY.
FT	DISULFID	1345	1358	BY SIMILARITY.
FT	DISULFID	1364	1377	BY SIMILARITY.
FT	DISULFID	1371	1386	BY SIMILARITY.
FT	DISULFID	1388	1399	BY SIMILARITY.
FT	DISULFID	1405	1418	BY SIMILARITY.
FT	DISULFID	1412	1427	BY SIMILARITY.
FT	DISULFID	1429	1440	BY SIMILARITY.
FT	DISULFID	1446	1458	BY SIMILARITY.
FT	DISULFID	1453	1467	BY SIMILARITY.
FT	DISULFID	1469	1482	BY SIMILARITY.
FT	DISULFID	1488	1499	BY SIMILARITY.
FT	DISULFID	1494	1508	BY SIMILARITY.
FT	DISULFID	1510	1523	BY SIMILARITY.
FT	DISULFID	1529	1540	BY SIMILARITY.
FT	DISULFID	1551	1549	BY SIMILARITY.
FT	DISULFID	1551	1564	BY SIMILARITY.
FT	DISULFID	1647	1659	BY SIMILARITY.
FT	DISULFID	1654	1668	BY SIMILARITY.
FT	DISULFID	1670	1683	BY SIMILARITY.
FT	DISULFID	1689	1701	BY SIMILARITY.
FT	DISULFID	1696	1710	BY SIMILARITY.
FT	DISULFID	1712	1725	BY SIMILARITY.
FT	DISULFID	1805	1817	BY SIMILARITY.
Query Match				11.2%;
Best Local Similarity				23.9%;
Matches 60; Conservative 24; Mismatches 83; Indels 84; Gaps 15;				
Oy	23	IPOCOANTCTPTNMMNN---HICRC---LAODEFMFSDDAGDST-----	63	
Db	2485	LDECSQSRPC--NICKNTKSGSYOCSPRGVLDGDKTKDDECKTKQKNCQFLCVN	2542	
Oy	64	-----DGF---HDICGPNKELDEE-----TCQ-----CVCRAGL---RP	91	
Db	2543	TUGFTCKCPGFTQHHRACIDNNECGSGPSTLCGAKGICQNTPGSFCECGRFSLSASG	2602	

```

OY 92 ASGPAPKELDRN-SCQVCNKKLFPSQCGA-----NREFDENTCQVCYKRTCPRN 140
DB 2603 LNEDEVDDECDGNHRCGHGONILGVRCCGPHGDVQHYQMNQCVDNES----- 2652
OY 141 OPLNPGKC---ACECTESPQKCLL-KGKFFHQTCSYR-RPCTNPKACEPGFSYSE-- 193
DB 2653 ---NPGACGASGAYNTLGSYKCAKCPGSPFSDQFSSACHDVNECSSKNPCSYGCSNTEGG 2709
OY 194 EYCRCPYSYK 204
DB 2710 YLCGCPGXYR 2720

RESULT 12
FBN2_HUMAN
ID FBN2_HUMAN STANDARD; PRT: 2911 AA.
AC P35556;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE FIBRILIN 2 PRECURSOR.
GN FBN2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-94165150; PubMed-8120105;
RA Zhang H., Apfelroth S.D., Hu W., Davis E.C., Sanguineti C.,
RA Bonadio J., Mecham R.P., Ramirez F.;
RT "Structure and expression of fibrillin-2, a novel microfibrillar
RT component preferentially located in elastic matrices.";
RL J. Cell Biol. 124:855-863(1994).
RN [2]
RP SEQUENCE OF 752-1505 FROM N.A.
RX MEDLINE-91304567; PubMed-1852206;
RA Lee B., Godfrey M., Vitale E., Hori H., Mattei M.-G., Sarfarazi M.,
RA Tsipouras P., Ramirez F., Hollister D.;
RT "Linkage of Marfan syndrome and a phenotypically related disorder to
RT two different fibrillin genes.";
RL Nature 352:330-334(1991).
RN [3]
RP VARIANTS CCA TYR-1252 AND SER-1433, AND VARIANT ILE-964.
RX MEDLINE-96083599; PubMed-7493032;
RA Putnam E.A., Zhang H., Ramirez F., Milewicz D.M.;
RT "Fibrillin-2 (FBN2) mutations result in the Marfan-like disorder,
RT congenital contractual arachnodactyly.";
RL Nat. Genet. 11:456-458(1995).
CC -1- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC THAT BINDS CALCIUM. FIBRILIN-2-CONTAINING MICROFIBRILS REGULATE
CC THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.
CC -1- DISEASE: DEFECTS IN FBN2 ARE THE CAUSE OF CONGENITAL CONTRACTURAL
CC ARACHNOIDACTYLY (CCA) (ALSO KNOWN AS BEALS SYNDROME). CCA IS
CC PHENOTYPICALLY SIMILAR TO MARFAN SYNDROME, BUT DOES NOT AFFECT THE
CC AORTA AND THE EYES.
CC -1- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
CC EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: U03272; AAA18950.1;
CC EMBL: X62009; -; NOT_ANNOTATED_CDS.
CC PIR: S17063; S17063.
CC PIR: S31101; S31101.

```

```

DR HSP; P35555; 1EMO.
DR MIM: 121050; -.
DR InterPro: IPR00152; -.
DR InterPro: IPR000561; -.
DR InterPro: IPR001438; -.
DR InterPro: IPR001881; -.
DR InterPro: IPR002212; -.
DR Pfam: PF00008; EGF; 46.
DR Pfam: PF00683; TB; 9.
DR PRINTS: PR00010; EGFBLD.
DR PROSITE: PS00010; ASX_HYDROXYL; 43.
DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS01186; EGF_2; 37.
DR PROSITE: PS01187; EGF_CA; 43.
DR Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW Repeat; Signal; Multigene family; Disease mutation; Polymorphism.
FT SIGNAL 1 28
FT CHAIN 29 2911
FT DOMAIN 111 142
FT DOMAIN 145 176
FT DOMAIN 176 207
FT DOMAIN 275 316
FT DOMAIN 317 358
FT REPEAT 359 425
FT DOMAIN 493 533
FT DOMAIN 534 573
FT DOMAIN 574 615
FT DOMAIN 616 656
FT REPEAT 657 766
FT DOMAIN 767 808
FT DOMAIN 809 850
FT DOMAIN 851 890
FT DOMAIN 954 995
FT REPEAT 996 1071
FT DOMAIN 1072 1113
FT DOMAIN 1114 1156
FT DOMAIN 1157 1198
FT DOMAIN 1199 1240
FT DOMAIN 1241 1281
FT DOMAIN 1282 1323
FT DOMAIN 1324 1365
FT DOMAIN 1366 1406
FT DOMAIN 1407 1447
FT DOMAIN 1448 1489
FT DOMAIN 1490 1530
FT DOMAIN 1531 1571
FT REPEAT 1572 1648
FT DOMAIN 1649 1690
FT REPEAT 1691 1732
FT REPEAT 1733 1806
FT DOMAIN 1807 1848
FT DOMAIN 1849 1890
FT DOMAIN 1891 1932
FT DOMAIN 1933 1971
FT DOMAIN 1972 2014
FT DOMAIN 2015 2054
FT DOMAIN 2055 2096
FT REPEAT 2097 2169
FT DOMAIN 2170 2211
FT DOMAIN 2212 2251
FT DOMAIN 2252 2292
FT DOMAIN 2293 2336
FT DOMAIN 2337 2378
FT REPEAT 2379 2447
FT DOMAIN 2448 2489
FT DOMAIN 2490 2530
FT DOMAIN 2531 2569
FT DOMAIN 2570 2612
FT DOMAIN 2613 2652
FT DOMAIN 2653 2694
FT DOMAIN 2733 2733
FT DISULFID 115 124

```

```

EGF-LIKE 1, NON-CALCIUM BINDING.
EGF-LIKE 2, NON-CALCIUM BINDING.
EGF-LIKE 3, NON-CALCIUM BINDING.
EGF-LIKE 4, CALCIUM-BINDING.
EGF-LIKE 5, CALCIUM-BINDING.
EGF-LIKE 6, NON-CALCIUM BINDING.
EGF-LIKE 7, CALCIUM-BINDING.
EGF-LIKE 8, CALCIUM-BINDING.
EGF-LIKE 9, CALCIUM-BINDING.
EGF-LIKE 10, CALCIUM-BINDING.
EGF-LIKE 11, CALCIUM-BINDING.
EGF-LIKE 12, CALCIUM-BINDING.
EGF-LIKE 13, CALCIUM-BINDING.
EGF-LIKE 14, CALCIUM-BINDING.
EGF-LIKE 15, CALCIUM-BINDING.
EGF-LIKE 16, CALCIUM-BINDING.
EGF-LIKE 17, CALCIUM-BINDING.
EGF-LIKE 18, CALCIUM-BINDING.
EGF-LIKE 19, CALCIUM-BINDING.
EGF-LIKE 20, CALCIUM-BINDING.
EGF-LIKE 21, CALCIUM-BINDING.
EGF-LIKE 22, CALCIUM-BINDING.
EGF-LIKE 23, CALCIUM-BINDING.
EGF-LIKE 24, CALCIUM-BINDING.
EGF-LIKE 25, CALCIUM-BINDING.
EGF-LIKE 26, CALCIUM-BINDING.
EGF-LIKE 27, CALCIUM-BINDING.
EGF-LIKE 28, CALCIUM-BINDING.
EGF-LIKE 29, CALCIUM-BINDING.
EGF-LIKE 30, CALCIUM-BINDING.
EGF-LIKE 31, CALCIUM-BINDING.
EGF-LIKE 32, CALCIUM-BINDING.
EGF-LIKE 33, CALCIUM-BINDING.
EGF-LIKE 34, CALCIUM-BINDING.
EGF-LIKE 35, CALCIUM-BINDING.
EGF-LIKE 36, CALCIUM-BINDING.
EGF-LIKE 37, CALCIUM-BINDING.
EGF-LIKE 38, CALCIUM-BINDING.
EGF-LIKE 39, CALCIUM-BINDING.
EGF-LIKE 40, CALCIUM-BINDING.
EGF-LIKE 41, CALCIUM-BINDING.
EGF-LIKE 42, CALCIUM-BINDING.
EGF-LIKE 43, CALCIUM-BINDING.
EGF-LIKE 44, CALCIUM-BINDING.
EGF-LIKE 45, CALCIUM-BINDING.
EGF-LIKE 46, CALCIUM-BINDING.
EGF-LIKE 47, CALCIUM-BINDING.
BY SIMILARITY.

```

FT	DISULFID	119	130	BY SIMILARITY.
FT	DISULFID	132	141	BY SIMILARITY.
FT	DISULFID	149	159	BY SIMILARITY.
FT	DISULFID	153	164	BY SIMILARITY.
FT	DISULFID	166	175	BY SIMILARITY.
FT	DISULFID	180	190	BY SIMILARITY.
FT	DISULFID	184	195	BY SIMILARITY.
FT	DISULFID	197	206	BY SIMILARITY.
FT	DISULFID	279	291	BY SIMILARITY.
FT	DISULFID	286	300	BY SIMILARITY.
FT	DISULFID	302	315	BY SIMILARITY.
FT	DISULFID	321	333	BY SIMILARITY.
FT	DISULFID	328	342	BY SIMILARITY.
FT	DISULFID	344	357	BY SIMILARITY.
FT	DISULFID	497	509	BY SIMILARITY.
FT	DISULFID	504	518	BY SIMILARITY.
FT	DISULFID	520	532	BY SIMILARITY.
FT	DISULFID	538	548	BY SIMILARITY.
FT	DISULFID	543	557	BY SIMILARITY.
FT	DISULFID	559	572	BY SIMILARITY.
FT	DISULFID	578	590	BY SIMILARITY.
FT	DISULFID	585	599	BY SIMILARITY.
FT	DISULFID	601	614	BY SIMILARITY.
FT	DISULFID	620	631	BY SIMILARITY.
FT	DISULFID	626	640	BY SIMILARITY.
FT	DISULFID	642	655	BY SIMILARITY.
FT	DISULFID	661	672	BY SIMILARITY.
FT	DISULFID	667	681	BY SIMILARITY.
FT	DISULFID	683	696	BY SIMILARITY.
FT	DISULFID	771	783	BY SIMILARITY.
FT	DISULFID	778	792	BY SIMILARITY.
FT	DISULFID	794	807	BY SIMILARITY.
FT	DISULFID	813	825	BY SIMILARITY.
FT	DISULFID	820	834	BY SIMILARITY.
FT	DISULFID	836	849	BY SIMILARITY.
FT	DISULFID	855	865	BY SIMILARITY.
FT	DISULFID	860	874	BY SIMILARITY.
FT	DISULFID	876	889	BY SIMILARITY.
FT	DISULFID	958	970	BY SIMILARITY.
FT	DISULFID	965	979	BY SIMILARITY.
FT	DISULFID	981	994	BY SIMILARITY.
FT	DISULFID	1076	1088	BY SIMILARITY.
FT	DISULFID	1083	1097	BY SIMILARITY.
FT	DISULFID	1099	1112	BY SIMILARITY.
FT	DISULFID	1118	1130	BY SIMILARITY.
FT	DISULFID	1125	1139	BY SIMILARITY.
FT	DISULFID	1141	1155	BY SIMILARITY.
FT	DISULFID	1161	1173	BY SIMILARITY.
FT	DISULFID	1168	1182	BY SIMILARITY.
FT	DISULFID	1184	1197	BY SIMILARITY.
FT	DISULFID	1203	1215	BY SIMILARITY.
FT	DISULFID	1210	1224	BY SIMILARITY.
FT	DISULFID	1226	1239	BY SIMILARITY.
FT	DISULFID	1245	1256	BY SIMILARITY.
FT	DISULFID	1267	1280	BY SIMILARITY.
FT	DISULFID	1286	1298	BY SIMILARITY.
FT	DISULFID	1293	1307	BY SIMILARITY.
FT	DISULFID	1309	1322	BY SIMILARITY.
FT	DISULFID	1328	1340	BY SIMILARITY.
FT	DISULFID	1335	1349	BY SIMILARITY.
FT	DISULFID	1351	1364	BY SIMILARITY.
FT	DISULFID	1370	1383	BY SIMILARITY.
FT	DISULFID	1377	1392	BY SIMILARITY.
FT	DISULFID	1394	1405	BY SIMILARITY.
FT	DISULFID	1411	1424	BY SIMILARITY.
FT	DISULFID	1418	1433	BY SIMILARITY.
FT	DISULFID	1435	1446	BY SIMILARITY.
FT	DISULFID	1452	1464	BY SIMILARITY.
FT	DISULFID	1459	1473	BY SIMILARITY.

Query Match 11.1%; Score 135; DB 1; length 2911;
Best Local Similarity 23.9%; Pred. No. 0.0007;

Matches	60;	Conservative	23;	Mismatches	84;	Indels	84;	Gaps	15;
Oy	23	LPQCAANTCFNFMNMMN----	HICRC----	LAODEMFSSDADDDT-----	63				
Db	2491	LDKCSQSPKPC--NYTCKNFBESYQSCPGYVLDGKTCXDLDECQKQHNCFLCVN	2548						
Oy	64	-----DGF-----HDICGPNKELDEE-----TCQ-----CVCRAGL-----RP	91						
Db	2549	TLGFTTCCKPFGFTQHTHTACIDNNECGSPILLCGKGICQNTPGSFSCGQGFSLDATTG	2608						
Oy	92	ASCGPHKEIDRN--SCQCVCKNKLFPSSQCA-----NEPDENTCQCVCKRTCPBN	140						
Db	2609	INCEVDDECDGNHRCQHNOCQNTLGGYRCGCPQGYIQHYQMNOCVDNECS-----	2658						
Oy	141	QPLNKGK--ACECTESPQKLL-KGKKFHQTSQCR-RCCTNRQKRCGEFGFSYSE--	193						
Db	2659	---NPNACGSASCYNTLGSYKCACPGFSFDQFSSACHDVNCCSSKNPCNVCNTEGG	2715						
Oy	194	EVCRCVPSYWK	204						
Db	2716	YLCGCPGYR	2726						
RESULT	13								
FN1_HUMAN									
ID	FN1_HUMAN	STANDARD:	PRT:	2871	AA.				
AC	P35555:								
DT	01-JUN-1994 (Rel. 29, Created)								
DT	01-JUN-1994 (Rel. 29, Last sequence update)								
DT	15-DEC-1998 (Rel. 37, Last annotation update)								
DE	FIBRILIN 1 PRECURSOR.								
GN	FN1 OR FN1.								
OS	Homo sapiens (Human).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.								
OX	NCBI_TaxID=9606;								
NP	[1]								
RC	SEQUENCE FROM N.A.								
RX	TISSUE=Placenta; PubMed=8364578;								
RA	MEDLINE=93372860;								
RA	Pereira L.V., D'Alessio M., Ramirez F., Lynch J.R., Sykes B.,								
RT	Pangilinan T., Bonadio J.;								
RT	"Genomic organization of the sequence coding for fibrillin, the								
RT	defective gene product in Marfan syndrome.";								
RL	Hum. Mol. Genet. 2:961-968(1993).								
NP	[2]								
RC	SEQUENCE OF 1-932 FROM N.A.								
RX	TISSUE=Placenta; and Fibroblast;								
RX	MEDLINE=94010947; PubMed=7691719;								
RA	Corson G.M., Chabberg S.C., Dietz H.C., Charbonneau N.L., Sakai L.Y.;								
RT	"Fibrillin binds calcium and is coded by cDNAs that reveal a								
RT	multidomain structure and alternatively spliced exons at the 5'								
RT	end.";								
RL	Genomics 17:476-484(1993).								
NP	[3]								
RC	SEQUENCE OF 899-2871 FROM N.A.								
RX	MEDLINE=91304568; PubMed=1852207;								
RA	Maslen C.L., Corson G.M., Maddox B.K., Glanville R.W., Sakai L.Y.;								
RT	"Partial sequence of a candidate gene for the Marfan syndrome.";								
RL	Nature 352:334-337(1991).								
NP	[4]								
RC	SEQUENCE OF 813-1313 FROM N.A.								
RX	MEDLINE=91304567; PubMed=1852206;								
RA	Lee B., Godfrey M., Vitale E., Hori H., Mattei M.-G., Sarfarazi M.,								
RA	Tsipouras P., Ramirez F., Hollister D.W.;								
RT	"Linkage of Marfan syndrome and a phenotypically related disorder to								
RT	two different fibrillin genes.";								
RL	Nature 352:330-334(1991).								
NP	[5]								
RC	CHARACTERIZATION.								
RX	MEDLINE=91317849; PubMed=1860873;								
RA	Sakai L.Y., Keene D.R., Glanville R.W., Bachinger H.P.;								
RT	"Purification and partial characterization of fibrillin, a								

RT cysteine-rich structural component of connective tissue
 RT microfibrils.";
 RN J. Biol. Chem. 266:14763-14770(1991).
 RN [6]
 RP STRUCTURE BY NMR OF 2054-2125.
 RX MEDLINE=98031893; PubMed=9362480;
 RA Yuan X., Downing A.K., Knott V., Handford P.A.;
 RT "Solution structure of the transforming growth factor beta-binding
 RT protein-like module, a domain associated with matrix fibrils.";
 RL EMBO J. 16:6659-6666(1997).
 RN [7]
 RP STRUCTURE BY NMR OF 2124-2205.
 RX MEDLINE=96144829; PubMed=8568869;
 RA Knott V., Downing A.K., Cardy C.M., Handford P.A.;
 RT "Calcium binding properties of an epidermal growth factor-like domain
 RT pair from human fibrillin-1.";
 RL J. Mol. Biol. 255:22-27(1996).
 RN [8]
 RP STRUCTURE BY NMR OF 2124-2205.
 RX MEDLINE=96222301; PubMed=8653794;
 RA Downing A.K., Knott V., Werner J.M., Cardy C.M., Campbell I.D.,
 RA Handford P.A.;
 RT "Solution structure of a pair of calcium-binding epidermal growth
 RT factor-like domains: implications for the Marfan syndrome and other
 RT genetic disorders.";
 RL Cell 85:597-605(1996).
 RN [9]
 RP REVIEW ON MFS VARIANTS.
 RX MEDLINE=96174615; PubMed=8594563;
 RA Collod G., Beroud C., Soussi T., Junien C., Boileau C.;
 RT "Software and database for the analysis of mutations in the human
 RT FBN1 gene.";
 RL Nucleic Acids Res. 24:137-141(1996).
 RN [10]
 RP REVIEW ON MFS VARIANTS.
 RX MEDLINE=97169383; PubMed=9016526;
 RA Collod-Beroud G., Beroud C., Ades L., Black C., Boxer M., Brock D.J.,
 RA Godfrey M., Hayward C., Karttunen L., Milewicz D., Peltonen L.,
 RA Richards R.I., Wang W., Junien C., Boileau C.;
 RT "Marfan Database (second edition): software and database for the
 RT analysis of mutations in the human FBN1 gene.";
 RL Nucleic Acids Res. 25:147-150(1997).
 RN [11]
 RP REVIEW ON VARIANTS.
 RX MEDLINE=98062175; PubMed=9401003;
 RA Hayward C., Brock D.J.H.;
 RT "Fibrillin-1 mutations in Marfan syndrome and other type-1
 RT fibrillogenopathies.";
 RL Hum. Mutat. 10:415-423(1997).
 RN [12]
 RP VARIANT MFS PRO-1137.
 RX MEDLINE=91304569; PubMed=1852208;
 RA Dietz H.C., Cutting G.R., Pyeritz R.E., Maslen C.L., Sakai L.Y.,
 RA Corson G.M., Puffenberger E.G., Hamosh A., Nanthakumar E.J.,
 RA Curristin S.M., Stetten G., Meyers D.A., Francomano C.A.;
 RT "Marfan syndrome caused by a recurrent de novo missense mutation in
 RT the fibrillin gene.";
 RL Nature 352:337-339(1991).
 RN [13]
 RP VARIANTS MFS SER-1249; ARG-1663; SER-2221 AND SER-2307.
 RX MEDLINE=92250834; PubMed=1301946;
 RA Dietz H.C., Saraiva J.M., Pyeritz R.E., Cutting G.R., Francomano C.A.;
 RT "Clustering of fibrillin (FBN1) missense mutations in Marfan syndrome
 RT patients at cysteine residues in EGF-like domains.";
 RL Hum. Mutat. 1:366-374(1992).
 RN [14]
 RP VARIANT MFS SER-2307.
 RX MEDLINE=92235290; PubMed=1569206;
 RA Dietz H.C., Pyeritz R.E., Puffenberger E.G., Kendzior R.J. Jr.,
 RA Corson G.M., Maslen C.L., Sakai L.Y., Francomano C.A., Cutting G.R.;
 RT "Marfan phenotype variability in a family segregating a missense
 RT mutation in the epidermal growth factor-like motif of the fibrillin
 RT gene.";
 RN J. Clin. Invest. 89:1674-1680(1992).
 RN [15]
 RP VARIANTS MFS ILE-548 AND ALA-723.
 RX MEDLINE=94010946; PubMed=8406497;
 RA Dietz H.C., McIntosh I., Sakai L.Y., Corson G.M., Chalberg S.C.,
 RA Pyeritz R.E., Francomano C.A.;
 RT "Four novel FBN1 mutations: significance for mutant transcript level
 RT and EGF-like domain calcium binding in the pathogenesis of Marfan
 RT syndrome.";
 RL Genomics 17:468-475(1993).
 RN [16]
 RP VARIANTS MFS SER-2144.
 RX MEDLINE=93278402; PubMed=8504310;
 RA Hewett D.R., Lynch J.R., Smith R., Sykes B.C.;
 RT "A novel fibrillin mutation in the Marfan syndrome which could
 RT disrupt calcium binding of the epidermal growth factor-like module.";
 RL Hum. Mol. Genet. 2:475-477(1993).
 RN [17]
 RP VARIANTS MFS R-862; Y-1117; P-1137 AND F-1589, AND VARIANT A-1148.
 RX MEDLINE=94108431; PubMed=8281141;
 RA Tynan K., Comeau K., Pearson M., Milgenbus P., Levitt D., Gasner C.,
 RA Berg M.A., Miller D.C., Francke U.;
 RT "Mutation screening of complete fibrillin-1 coding sequence: report
 RT of five new mutations, including two in 8-cysteine domains.";
 RL Hum. Mol. Genet. 2:1813-1821(1993).
 RN [18]
 RP VARIANTS MFS GLY-217 AND ARG-2627.
 RX MEDLINE=95067970; PubMed=7977366;
 RA Karttunen L., Raghunath M., Loeonqvist L., Peltonen L.;
 RT "A compound-heterozygous Marfan patient: two defective fibrillin
 RT alleles result in a lethal phenotype.";
 RL Am. J. Hum. Genet. 55:1083-1091(1994).
 RN [19]
 RP VARIANT EL IYS-2447.
 RX MEDLINE=94245249; PubMed=8186302;
 RA Lonnqvist L., Child A., Kainulainen K., Davidson R., Puhakka L.,
 RA Peltonen L.;
 RT "A novel mutation of the fibrillin gene causing ectopia lentis.";
 RL Genomics 19:573-576(1994).
 RN [20]
 RP VARIANT MFS CYS-627.
 RX MEDLINE=94272487; PubMed=8004112;
 RA Hayward C., Rae A.L., Porteous M.E.M., Logie L.J., Brock L.J.;
 RT "Two novel mutations and a neutral polymorphism in EGF-like domains
 RT of the fibrillin gene (FBN1): SSCP screening of exons 15-21 in Marfan
 RT syndrome patients.";
 RL Hum. Mol. Genet. 3:373-375(1994).
 RN [21]
 RP VARIANT MFS CYS-122.
 RX MEDLINE=94314977; PubMed=8040326;
 RA Stahl-Hallengren C., Orkonen T., Kainulainen K., Kristofersson U.,
 RA Saxne T., Tornqvist K., Peltonen L.;
 RT "An extra cysteine in one of the non-calcium-binding epidermal growth
 RT factor-like motifs of the FBN1 polypeptide is connected to a novel
 RT variant of Marfan syndrome.";
 RL J. Clin. Invest. 94:709-713(1994).
 RN [22]
 RP VARIANT MFS TYR-1223.
 RX MEDLINE=94351682; PubMed=8071963;
 RA Hewett D.R., Lynch J.R., Child A., Sykes B.C.;
 RT "A new missense mutation of fibrillin in a patient with Marfan
 RT syndrome.";
 RL J. Med. Genet. 31:338-339(1994).
 RN [23]
 RP VARIANT MFS HIS-1170.
 RX MEDLINE=95174777; PubMed=7870075;
 RA Hayward C., Porteous M.E.M., Brock D.J.H.;
 RT "A novel mutation in the fibrillin gene (FBN1) in familial
 RT archonodactyly.";
 RL Mol. Cell. Probes 8:325-327(1994).
 RN [24]
 RP VARIANTS MFS G217; N1023; R1074; Y1242; R1513; E2127; W2151; R2447 AND R2511.
 RX MEDLINE=94184368; PubMed=8136837;

RA Kainulainen K., Karttunen L., Puhakka L., Sakai L., Peltonen L.;
 RT "Mutations in the fibrillin gene responsible for dominant ectopia
 RL lens and neonatal Marfan syndrome."
 RL Nat. Genet. 6:64-69(1994).
 RL [25]
 RP VARIANT SER-1127.

Query Match 10.8%; Score 131.5; DB 1; Length 2871;
 Best Local Similarity 19.0%; Pred. No. 0.0014;
 Matches 56; Conservative 32; Mismatches 89; Indels 117; Gaps 15;

QY 5 LDYROYHSTIRSLPATLPOCO-----AANKTCPPNYMNNHICRCLA 48
 DB 1757 IDLY-----TGLPYVDIDCEIRPGVCEYCYNNMVGSEFCECVGFYNDKLLVC-- 1806
 QY 49 QEDFMESDAGDSTGDFHIDCGPNKL-----DEEFCQ-- 82
 DB 1807 -EDI-----DECQNG--PVCGRNAECINTAGSYRCDCPKGYRTSTGQCNDRNECQET 1856
 QY 83 -----CVCRAGLR-----PASCGRHKEIDLRNSC-OCVCKNKLFPSCCGA 120
 DB 1857 PNCISGQCIDTVGSFYCLCHTGFKNIDQTMCLDINECERDAGCGTGNITGSENCRC 1916
 QY 121 NREF-----DENTCQCVCCKRTCPRMQPLNP-GKCAECTESPQ-----K 158
 DB 1917 NHGFIILSHNNDICIDVDECAAGNGLCRNGQCIWVGSPQOCNCEGYEVAADGRTCYDINE 1976
 QY 159 CLKGKGFHHQTC-----SCYRRPCTNRKAKCEGFSEYSEVCCVSYNKRPM 208
 DB 1977 CLLEPRKCAPGTQNLJDGSYR--C-----ICPPGYSLONKCRDIDECVEEPEI 2023

RESULT 14

FBN1_MOUSE STANDARD: PRT: 2871 AA.

ID FBN1_MOUSE
 AC 061354; Q60826;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE FIBRILLIN 1 PRECURSOR.
 GN FBN1 OR FBN-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95130561; PubMed=7829516;
 RA Yin W., Germiller J., Sanguinetti C., Smiley E., Pangillan T.,
 RA Pereira L., Ramirez F., Bonadio J.;
 RT "Primary structure and developmental expression of Fbn-1, the mouse
 RT fibrillin gene."
 RL J. Biol. Chem. 270:1798-1806(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD-1; TISSUE=Kidney;
 RA Ota K., Kumar A., Mada J., Liu Z., Kanwar Y.S.;
 RL Submitted (Apr-1995) to the EMBL/Genbank/DBP databases.
 CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
 CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS
 CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
 CC -!- PTM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER
 CC FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE
 CC MICROFIBRILS (BY SIMILARITY).
 CC -!- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
 CC EGF-LIKE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: U29454; AAA56840.1; -
 DR EMBL: U22493; AAA64217.1; -
 CC HSSP: P35355; IAPJ.
 DR MGI: 95489; Fbn1.
 DR InterPro: IPR000152; -
 DR InterPro: IPR000561; -
 DR InterPro: IPR001881; -
 DR InterPro: IPR002212; -
 DR Pfam: PF00008; EGF; 46.
 DR Pfam: PF00683; TB; 9.
 DR PROSITE: PS00010; ASX_HYDROXYL; 43.
 DR PROSITE: PS00022; EGF_1; 2.
 DR PROSITE: PS01186; EGF_2; 38.
 DR PROSITE: PS01187; EGF_CA; 45.
 KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
 KW Repeat; Signal; Multigene family.
 FT SIGNAL 1 27
 FT CHAIN 28 2871
 FT DOMAIN 81 112
 FT DOMAIN 115 146
 FT DOMAIN 147 178
 FT DOMAIN 246 287
 FT DOMAIN 288 329
 FT DOMAIN 330 401
 FT DOMAIN 402 446
 FT DOMAIN 449 489
 FT DOMAIN 490 529
 FT DOMAIN 530 571
 FT DOMAIN 572 612
 FT DOMAIN 613 653
 FT DOMAIN 656 721
 FT DOMAIN 723 764
 FT DOMAIN 765 806
 FT DOMAIN 807 846
 FT DOMAIN 910 951
 FT DOMAIN 952 1018
 FT DOMAIN 1028 1069
 FT DOMAIN 1070 1112
 FT DOMAIN 1113 1154
 FT DOMAIN 1155 1196
 FT DOMAIN 1197 1237
 FT DOMAIN 1238 1279
 FT DOMAIN 1280 1321
 FT DOMAIN 1322 1362
 FT DOMAIN 1363 1403
 FT DOMAIN 1404 1445
 FT DOMAIN 1446 1486
 FT DOMAIN 1487 1527
 FT DOMAIN 1528 1569
 FT DOMAIN 1569 1606
 FT DOMAIN 1606 1647
 FT DOMAIN 1648 1688
 FT DOMAIN 1689 1758
 FT DOMAIN 1766 1807
 FT DOMAIN 1808 1848
 FT DOMAIN 1849 1890
 FT DOMAIN 1891 1929
 FT DOMAIN 1930 1972
 FT DOMAIN 1973 2012
 FT DOMAIN 2013 2054
 FT DOMAIN 2055 2121
 FT DOMAIN 2127 2165
 FT DOMAIN 2166 2205
 FT DOMAIN 2206 2246
 FT DOMAIN 2247 2290
 FT DOMAIN 2291 2332
 FT DOMAIN 2333 2400
 FT DOMAIN 2402 2443
 FT DOMAIN 2444 2484
 FT DOMAIN 2485 2523
 FT DOMAIN 2524 2566
 FIBRILLIN 1.
 EGF-LIKE 1, NON-CALCIUM BINDING.
 EGF-LIKE 2, NON-CALCIUM BINDING.
 EGF-LIKE 3, NON-CALCIUM BINDING.
 EGF-LIKE 4, CALCIUM-BINDING.
 EGF-LIKE 5, CALCIUM-BINDING.
 TGF-
 PRO-RICH.
 EGF-LIKE 6, NON-CALCIUM BINDING.
 EGF-LIKE 7, CALCIUM-BINDING.
 EGF-LIKE 8, CALCIUM-BINDING.
 EGF-LIKE 9, CALCIUM-BINDING.
 EGF-LIKE 10, CALCIUM-BINDING.
 TGF-
 EGF-LIKE 11, CALCIUM-BINDING.
 EGF-LIKE 12, CALCIUM-BINDING.
 EGF-LIKE 13, CALCIUM-BINDING.
 EGF-LIKE 14, CALCIUM-BINDING.
 TGF-
 EGF-LIKE 15, CALCIUM-BINDING.
 EGF-LIKE 16, CALCIUM-BINDING.
 EGF-LIKE 17, CALCIUM-BINDING.
 EGF-LIKE 18, CALCIUM-BINDING.
 EGF-LIKE 19, CALCIUM-BINDING.
 EGF-LIKE 20, CALCIUM-BINDING.
 EGF-LIKE 21, CALCIUM-BINDING.
 EGF-LIKE 22, CALCIUM-BINDING.
 EGF-LIKE 23, CALCIUM-BINDING.
 EGF-LIKE 24, CALCIUM-BINDING.
 EGF-LIKE 25, CALCIUM-BINDING.
 EGF-LIKE 26, CALCIUM-BINDING.
 TGF-
 EGF-LIKE 27, CALCIUM-BINDING.
 EGF-LIKE 28, CALCIUM-BINDING.
 TGF-
 EGF-LIKE 29, CALCIUM-BINDING.
 EGF-LIKE 30, CALCIUM-BINDING.
 EGF-LIKE 31, CALCIUM-BINDING.
 EGF-LIKE 32, CALCIUM-BINDING.
 EGF-LIKE 33, CALCIUM-BINDING.
 EGF-LIKE 34, CALCIUM-BINDING.
 EGF-LIKE 35, CALCIUM-BINDING.
 TGF-
 EGF-LIKE 36, CALCIUM-BINDING.
 EGF-LIKE 37, CALCIUM-BINDING.
 EGF-LIKE 38, CALCIUM-BINDING.
 EGF-LIKE 39, CALCIUM-BINDING.
 EGF-LIKE 40, CALCIUM-BINDING.
 TGF-
 EGF-LIKE 41, CALCIUM-BINDING.
 EGF-LIKE 42, CALCIUM-BINDING.
 EGF-LIKE 43, CALCIUM-BINDING.
 EGF-LIKE 44, CALCIUM-BINDING.

FT	DOMAIN	2567	2606	EGF-LIKE 45, CALCIUM-BINDING.
FT	DOMAIN	2567	2606	EGF-LIKE 45, CALCIUM-BINDING.
FT	DOMAIN	2607	2647	EGF-LIKE 46, CALCIUM-BINDING.
FT	DOMAIN	2647	2687	EGF-LIKE 47, CALCIUM-BINDING.
FT	DISULFID	85	94	BY SIMILARITY.
FT	DISULFID	89	100	BY SIMILARITY.
FT	DISULFID	102	111	BY SIMILARITY.
FT	DISULFID	119	129	BY SIMILARITY.
FT	DISULFID	123	134	BY SIMILARITY.
FT	DISULFID	136	145	BY SIMILARITY.
FT	DISULFID	150	160	BY SIMILARITY.
FT	DISULFID	154	166	BY SIMILARITY.
FT	DISULFID	168	177	BY SIMILARITY.
FT	DISULFID	250	262	BY SIMILARITY.
FT	DISULFID	257	271	BY SIMILARITY.
FT	DISULFID	273	286	BY SIMILARITY.
FT	DISULFID	292	304	BY SIMILARITY.
FT	DISULFID	299	313	BY SIMILARITY.
FT	DISULFID	315	328	BY SIMILARITY.
FT	DISULFID	453	465	BY SIMILARITY.
FT	DISULFID	460	474	BY SIMILARITY.
FT	DISULFID	476	488	BY SIMILARITY.
FT	DISULFID	494	504	BY SIMILARITY.
FT	DISULFID	499	513	BY SIMILARITY.
FT	DISULFID	515	528	BY SIMILARITY.
FT	DISULFID	534	546	BY SIMILARITY.
FT	DISULFID	541	555	BY SIMILARITY.
FT	DISULFID	557	570	BY SIMILARITY.
FT	DISULFID	576	587	BY SIMILARITY.
FT	DISULFID	582	596	BY SIMILARITY.
FT	DISULFID	598	611	BY SIMILARITY.
FT	DISULFID	617	628	BY SIMILARITY.
FT	DISULFID	623	637	BY SIMILARITY.
FT	DISULFID	639	652	BY SIMILARITY.
FT	DISULFID	727	739	BY SIMILARITY.
FT	DISULFID	734	748	BY SIMILARITY.
FT	DISULFID	750	763	BY SIMILARITY.
FT	DISULFID	769	781	BY SIMILARITY.
FT	DISULFID	776	790	BY SIMILARITY.
FT	DISULFID	792	805	BY SIMILARITY.
FT	DISULFID	811	821	BY SIMILARITY.
FT	DISULFID	816	830	BY SIMILARITY.
FT	DISULFID	832	845	BY SIMILARITY.
FT	DISULFID	914	926	BY SIMILARITY.
FT	DISULFID	921	935	BY SIMILARITY.
FT	DISULFID	937	950	BY SIMILARITY.
FT	DISULFID	1032	1044	BY SIMILARITY.
FT	DISULFID	1039	1053	BY SIMILARITY.
FT	DISULFID	1055	1068	BY SIMILARITY.
FT	DISULFID	1074	1086	BY SIMILARITY.
FT	DISULFID	1081	1095	BY SIMILARITY.
FT	DISULFID	1097	1111	BY SIMILARITY.
FT	DISULFID	1117	1129	BY SIMILARITY.
FT	DISULFID	1124	1138	BY SIMILARITY.
FT	DISULFID	1140	1153	BY SIMILARITY.
FT	DISULFID	1159	1171	BY SIMILARITY.
FT	DISULFID	1166	1180	BY SIMILARITY.
FT	DISULFID	1182	1195	BY SIMILARITY.
FT	DISULFID	1201	1212	BY SIMILARITY.
FT	DISULFID	1208	1221	BY SIMILARITY.
FT	DISULFID	1223	1236	BY SIMILARITY.
FT	DISULFID	1242	1254	BY SIMILARITY.
FT	DISULFID	1249	1263	BY SIMILARITY.
FT	DISULFID	1265	1278	BY SIMILARITY.
FT	DISULFID	1284	1296	BY SIMILARITY.
FT	DISULFID	1291	1305	BY SIMILARITY.
FT	DISULFID	1307	1320	BY SIMILARITY.
FT	DISULFID	1326	1339	BY SIMILARITY.
FT	DISULFID	1333	1348	BY SIMILARITY.
FT	DISULFID	1350	1361	BY SIMILARITY.
FT	DISULFID	1367	1380	BY SIMILARITY.
FT	DISULFID	1374	1389	BY SIMILARITY.
FT	DISULFID	1391	1402	BY SIMILARITY.
FT	DISULFID	1408	1420	BY SIMILARITY.

FT	DISULFID	1415	1429	BY SIMILARITY.
FT	DISULFID	1431	1444	BY SIMILARITY.
FT	DISULFID	1450	1461	BY SIMILARITY.
FT	DISULFID	1456	1470	BY SIMILARITY.
FT	DISULFID	1472	1485	BY SIMILARITY.
FT	DISULFID	1491	1502	BY SIMILARITY.
FT	DISULFID	1497	1511	BY SIMILARITY.
FT	DISULFID	1513	1526	BY SIMILARITY.
FT	DISULFID	1610	1622	BY SIMILARITY.
FT	DISULFID	1617	1631	BY SIMILARITY.
FT	DISULFID	1633	1646	BY SIMILARITY.
FT	DISULFID	1652	1663	BY SIMILARITY.
FT	DISULFID	1658	1672	BY SIMILARITY.
FT	DISULFID	1674	1687	BY SIMILARITY.
FT	DISULFID	1770	1782	BY SIMILARITY.

Query Match 10.8%; Score 131.5; DB 1; Length 2871;
 Best Local Similarity 19.7%; Pred. No. 0.0014;
 Matches 62; Conservative 30; Mismatches 84; Indels 139; Gaps 19;

OY 5 LDVYRQVHSIIRSLPATLPOCO-----AANKTCPTNYMMNNHICRCLA 48
 DB 1757 IDLY-----TGLPVDIDDEKREITPGVCEGNCVIMNGSFRECEVGFYNDKLLVC-- 1806
 OY 49 QEDFMFSSDAGDSDTDG-----FHDTGPNKEL-----DEETCO----- 82
 DB 1807 -EDI-----DECONGPCLRNACINTAGSYRCDCCKGGRYLTSTGQCNDRNECEIPN 1858
 OY 83 -----CVCNAGLR-----PASCGRHKLDELNSC--QCVCCKKKLFPSSCGARR 122
 DB 1859 ICSHGQCIDTVGSFYCLCHTGFETNEDQTMCLDINECDACNGTCRNMTISFNCRCNH 1918
 OY 123 EF-----DE-----NTC-----QCVCCK-----RRCP--RNOP 142
 DB 1919 GFTLSHNNCICVDDECATNGNLCNRNGQCVNTYGSQCNBSGYEAPRGRTICVDINECV 1978
 OY 143 LNPGRCA---CECTESPQCLL-KGKKFHHQC---SCYRRP-----CTNRQKA--- 184
 DB 1979 LDPGKCAPTCQMLDGSYCLCPGYSLONDKCEDIDEVEEPEICALGTSWTEGSPFC 2038
 OY 185 -CEPGFSYEVCRC 198
 DB 2039 LCEPGFSWSSSGRR 2053

RESULT 15
 EMRL_MOUSE
 ID EMRL_MOUSE STANDARD; PRT: 931 AA.
 AC Q61549;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update).
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CELL SURFACE GLYCOPROTEIN EMRI PRECURSOR (EMRI HORMONE RECEPTOR)
 DE (CELL SURFACE GLYCOPROTEIN F4/80).
 GN EMRI OR GPF480.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BAIB/C; TISSUE=peritoneal cavity;
 RX MEDLINE=96132946; PubMed=8550607;
 RA McKnight A.J., Macfarlane A.J., Drl P., Turley L., Willis A.C.,
 Gordon S.;
 RT "Molecular cloning of F4/80, a murine macrophage-restricted cell
 RT surface glycoprotein with homology to the G-protein-linked
 RT transmembrane 7 hormone receptor family.";
 RL J. Biol. Chem. 271:486-489(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97312684; PubMed=9169125;
 RA Lin H.H., Stubbs L.J., Mucenski M.L.;

Identification and characterization of a seven transmembrane hormone receptor using differential display.;
 Genomics 41:301-308(1997).
 - FUNCTION: PROBABLY INVOLVED IN CELL ADHESION WITHIN TISSUES
 AND RECEPTOR SIGNALING.
 - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 - TISSUE SPECIFICITY: IN MACROPHAGES; BUT ABSENT FROM THOSE WHICH ARE LOCALIZED WITHIN T-CELL AREAS OF LYMPH NODES AND SPLEEN.
 LOW LEVEL OF EXPRESSION ON BLOOD MONOCYTES.
 - SIMILARITY: CONTAINS 7 EGF-LIKE DOMAINS.
 - SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).

 CC EMBL: X93328; CA63720.1;
 CC EMBL: U66888; AAC53184.1;
 CC HSP: P07204; IFGD.
 DR GCRDb: GCR_1309;
 DR MGD: MGI:106912; Emr1.
 DR InterPro: IPR000152;
 DR InterPro: IPR000203;
 DR InterPro: IPR000561;
 DR InterPro: IPR000832;
 DR InterPro: IPR001740;
 DR InterPro: IPR001881;
 DR Pfam: PF00002; 7tm_2; 1.
 DR Pfam: PF00008; EGF; 7.
 DR Pfam: PF01825; GRS; 1.
 DR PRINTS: PRO1128; EMR1HOMONER.
 DR PROSITE: PS00650; G_PROTEIN_RECP_F2_2; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 6.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS01187; EGF_CA; 5.
 DR G-protein coupled receptor; Transmembrane; Receptor; Glycoprotein;
 KM EGF-like domain; Repeat; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 931
 FT DOMAIN 28 644
 FT TRANSMEM 645 672
 FT DOMAIN 673 679
 FT TRANSMEM 680 701
 FT DOMAIN 702 711
 FT TRANSMEM 712 735
 FT DOMAIN 736 754
 FT TRANSMEM 755 776
 FT DOMAIN 777 792
 FT TRANSMEM 793 821
 FT DOMAIN 822 839
 FT TRANSMEM 840 859
 FT DOMAIN 860 874
 FT TRANSMEM 875 897
 FT DOMAIN 898 931
 FT DOMAIN 932 80
 FT DOMAIN 81 132
 FT DOMAIN 133 172
 FT DOMAIN 173 221
 FT DOMAIN 222 271
 FT DOMAIN 272 318
 FT DOMAIN 319 367
 FT SITE 506 508
 FT DISULFID 36 48
 FT DISULFID 42 57
 FT DISULFID 59 79
 FT DISULFID 85 98
 FT DISULFID 92 107
 FT DISULFID 109 131
 FT DISULFID 137 149

FT DISULFID 143 158 BY SIMILARITY.
 FT DISULFID 160 171 BY SIMILARITY.
 FT DISULFID 177 189 BY SIMILARITY.
 FT DISULFID 183 198 BY SIMILARITY.
 FT DISULFID 200 220 BY SIMILARITY.
 FT DISULFID 226 239 BY SIMILARITY.
 FT DISULFID 233 248 BY SIMILARITY.
 FT DISULFID 250 270 BY SIMILARITY.
 FT DISULFID 276 286 BY SIMILARITY.
 FT DISULFID 280 295 BY SIMILARITY.
 FT DISULFID 297 317 BY SIMILARITY.
 FT DISULFID 323 336 BY SIMILARITY.
 FT DISULFID 330 345 BY SIMILARITY.
 FT DISULFID 347 366 BY SIMILARITY.
 FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 167 167 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 229 229 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 283 283 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 417 417 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 474 474 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 498 498 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 706 706 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 931 AA; 102129 MW; 52963A67EBB76B5 CRC64;

Query Match 10.7%; Score 130; DB 1; Length 931;
 Best Local Similarity 21.6%; Pred. No. 0.00074;
 Matches 65; Conservative 18; Mismatches 76; Indels 142; Gaps 16;

QY .25 OCGANKTCPTVMNNH-----CRLAQEDFMFSSDAGD-----STGEF-HDI----- 69
 DB 84 ECLGSDSPCGPNSVCTNIGRAKSCLRG-----FSSSTGKMDITGLSLDFLCAVDDECIT 139
 QY 70 -----CGPNKEL-----DEETC-----OCV 84
 DB 140 IGICPKYSCNSVGSYSCTCGPFLVNGSIDEDDECYTRVCPENAHNTLTLSYCT 199
 QY 85 CRAGLRPAPSCGP-----KELDRNSQCQ-----VCKKKLPSSCGANRREDENTC 129
 DB 200 CNSGLESSGGGPMFOGLDESCEDVEDCSRNSYLCGPTFCINTLTGSSCSCPAGFSLPF 259
 QY 130 OCY-----CKRTCPRN-----QPLNPGKAC----- 150
 DB 260 QTLGHPADGNCCTDIDCDPTCPLNSSCTNTIGSYCTCHPGFASSNGQLNFKOLEVTECD 319
 QY 151 --ECTESPQKCLLKGRKFFHQTSCYRRPCTNRQKACERGFSSSEVRCVPSYKRPQM 208
 DB 320 IDECTQDPLQGLNS-----VCIN-----VPG-SY-----IGCLDPPQMDPEBG 357
 QY 209 S 209
 DB 358 S 358

Search completed: October 17, 2001, 14:51.43
 Job time: 469 sec

Db 272 DSADGFHDICGNKELDEETCCVCCKGLAASSCGPHKELDRDSCQVCYCKNKLFPSSCGA 331
 QY 121 NREPDNTCCQCKRTCPNPNPLNPKGCACECTESPQCKLLGKKRHHOTCSYRRPCTN 180
 Db 332 NREPDNTCCQCKRTCPNPNPLNPKGCACECTENPKCKFLGKKRHHOTCSYRRPCTN 391
 QY 181 ROKACEGFSYSEVQCRVSYWKRPM 209
 Db 392 RYKHCEGGLSFSEVQCRVSYWKRPM 420

RESULT 2
 ID 057352 PRELIMINARY; PRT: 418 AA.
 AC 057352;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR.
 GN VEGF-C.
 OS Colurnix colurnix japonica (Japanese quail).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Colurnix.
 OC NCBI_TaxID=93934;
 RX MEDLINE=9435294;
 RA Eichmann A., Corbel C., Jaffredo T., Breant V., Joukov V., Kumar V.,
 Alltalo K., Le Douarin N.M.;
 "Avian VEGF-C: cloning, embryonic expression pattern and stimulation
 of the differentiation of VEGFR2-expressing endothelial cell
 precursors.";
 RT Development 125:743-752(1998).
 RL EMBL: Y15837; CAA75799.1; -.
 DR HSSP; P15692; 1VPF.
 DR InterPro; IPR000072; -.
 DR InterPro; IPR002400; -.
 DR Pfam; PF00341; PDGF_1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR Prodom; PD001629; -; 1.
 DR PROSITE; PS00249; PDGF_1; 1.
 DR PROSITE; PS50278; PDGF_2; 1.
 DR SMART; SM00141; PDGF; 1.
 DR SIGNAL.
 KW signal.
 FT CHAIN 1 31 POTENTIAL.
 FT CHAIN 111 418 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
 FT SIGNAL 111 418
 FT SEQUENCE 418 AA; 46839 MW; 099BFCC79151BF2B CRC64;

Query Match 76.1%; Score 927; DB 13; Length 418;
 Best Local Similarity 75.2%; Pred. No. 6.7e-93;
 Matches 158; Conservative 17; Mismatches 33; Indels 2; Gaps 2;

QY 1 CMSKLDVYRQVHSIIRSLPATLPCCQANKTCPTNYMNNHICCLAQEDPMFSSDAGD 60
 Db 210 CMSKLDVYRQVHSIIRSLPATLPCCQANKTCPTNYMNNHICCLAQEDPMFSSDAGD 269
 QY 61 -DSTGDFHDICGNKELDEETCCVCCKGLAASSCGPHKELDRDSCQVCYCKNKLFPSSCG 119
 Db 270 DSTGDFHDICGNKELDEETCCVCCKGLAASSCGPHKELDRDSCQVCYCKNKLFPSSCG 328
 QY 120 ANREPDNTCCQCKRTCPNPNPLNPKGCACECTESPQCKLLGKKRHHOTCSYRRPCT 179
 Db 329 ANREPDNTCCQCKRTCPNPNPLNPKGCACECTENPKCKFLGKKRHHOTCSYRRPCT 388
 QY 180 ROKACEGFSYSEVQCRVSYWKRPM 209
 Db 389 ROKACEGFSYSEVQCRVSYWKRPM 418

RESULT 3
 035757

ID 035757 PRELIMINARY; PRT: 126 AA.
 AC 035757;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR C (FRAGMENT).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RX MEDLINE=9435294;
 RA Mandriota S.J., Pepper M.S.;
 "Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases."
 DR EMBL; AF010302; AAB63248.1; -.
 DR HSSP; P15692; 2VPF.
 DR InterPro; IPR000072; -.
 DR Pfam; PF00341; PDGF; 1.
 DR PROSITE; PS50278; PDGF_2; 1.
 DR SMART; SM00141; PDGF; 1.
 FT NON_TER 1 126
 FT NON_TER 126 126
 FT SEQUENCE 126 AA; 13977 MW; 8F365AFBC4E037B0 CRC64;

Query Match 28.3%; Score 345; DB 11; Length 126;
 Best Local Similarity 80.6%; Pred. No. 4.5e-30;
 Matches 58; Conservative 10; Mismatches 4; Indels 0; Gaps 0;

QY 1 CMSKLDVYRQVHSIIRSLPATLPCCQANKTCPTNYMNNHICCLAQEDPMFSSDAGD 60
 Db 55 CMSKLDVYRQVHSIIRSLPATLPCCQANKTCPTNYMNNHICCLAQEDPMFSSDAGD 114
 QY 61 -DSTGDFHDICGP 72
 Db 115 -DSTGDFHDICGP 126

RESULT 4
 ID 043915 PRELIMINARY; PRT: 354 AA.
 AC 043915;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE GROWTH FACTOR FIGF.
 GN FIGF OR VEGF-D.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RX MEDLINE=9479493;
 RA Rochegiani M., Lestingi M., Luddi A., Orlandini M., Franco B.,
 Rossi E., Ballabio A., Zuffardi O., Oliviero S.;
 "Human FIGF: cloning, gene structure, and mapping to chromosome Xp22.1
 between the FIGA and the GPR genes.";
 RT Genomics 47:207-216(1998).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP TISSUE=LUNG;
 RC MEDLINE=97349118; PubMed=9205122;
 RA Yamada Y., Nezu J., Shimane M., Hirata Y.;
 "Molecular cloning of a novel vascular endothelial growth factor,
 VEGF-D.";
 RT Genomics 42:483-488(1997).
 RL [3]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=98118549; PubMed=9435229;
 RA Achen K., Jeltsch W., Kuk E., Maekinen T., Vitelli A., Wilks A.F.,
 Alltalo K., Stacker S.A.;

RT "Vascular endothelial growth factor D (VEGF-D) is a ligand for the
RT tyrosine kinases VEGF receptor 2 (Flk1) and VEGF receptor 3 (Flt4).";
RL Proc. Natl. Acad. Sci. U.S.A. 95:548-553(1998).
DR EMBL; Y12864; CAA73371.1; -. JOINED.
DR EMBL; Y12865; CAA73371.1; JOINED.
DR EMBL; Y12866; CAA73371.1; JOINED.
DR EMBL; Y12867; CAA73371.1; JOINED.
DR EMBL; Y12868; CAA73371.1; JOINED.
DR EMBL; Y12869; CAA73371.1; JOINED.
DR EMBL; Y12870; CAA73371.1; JOINED.
DR EMBL; D89630; BA24264.1; -. JOINED.
DR EMBL; AJ000185; CAA03942.1; -. JOINED.
DR EMBL; Y12863; CAA73370.1; -. JOINED.
DR HSSP; P15692; IVP.
DR InterPro; IPR000072; -.
DR Pfam; PF00341; PDGF_1.
DR ProDom; PD001629; -.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS0278; PDGF_2; 1.
DR SMART; SM00141; PDGF; 1.
SQ SEQUENCE 354 AA; 40444 MW; 20480769D735173E CRC64;

Query Match 20.8%; Score 253.5; DB 4; Length 354;
Best Local Similarity 30.1%; Pred. No. 9.6e-20;
Matches 59; Conservative 27; Mismatches 49; Indels 61; Gaps 8;

OY 9 ROVHSIIRSLPA-TLPQCAANKTCPTNYMNNHICRCLAQEDFMFSSDAGDSDTG 65
DB 197 RHPYSIIRSLT-QIPEDRCSHSKLCPIDMLDMSNCKKCLADE-----ENPLAG 245
OY 66 FHDICGPNKLEDETCQVCVRAGLRPASCGRPKELDRNSCCVCCKNKLFPSCGANREFD 125
DB 246 TTD-----HSHLOE-----PALCGPH-----MMFD 265
OY 126 ENTQCQVCCKRTCPNQPPLNPGKAC-ECTESPQCLLKGKFNHQTSC-----YRRC 178
DB 266 EDRCCVCCKTPCPKDLIQPKNCKSECEKESLETCCQKHKLFHPRTSCEDRCPTHTRC 325
OY 179 TNRQKACRPGFSYSEE 194
DB 326 ASGKTACAKHCRPKR 341

RESULT 5
O94446 PRELIMINARY; PRT; 1704 AA.
AC O94446;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE 220 KDA SILK PROTEIN.
GN SP220.
OS Chironomus thummi thummi (Midge).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomidae; Chironomidae; Chironominae; Chironomus.
OX NCBI_TaxID=7155;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-SALIVARY GLAND;
RA Case S.T., Cox C., Bell W.C., Hoffman R.T., Martin J., Hamilton R.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U54641; AAA99804.1; -.
DR InterPro; IPR000561; -.
DR InterPro; IPR000853; -.
DR PRINTS; PR00876; MTNEMATODE.
DR PROSITE; PS00222; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
SQ SEQUENCE 1704 AA; 185745 MW; 3A3F20247C8F1E28 CRC64;

Query Match 20.2%; Score 245.5; DB 5; Length 1704;

Best Local Similarity 27.6%; Pred. No. 2.7e-18;
Matches 56; Conservative 31; Mismatches 75; Indels 41; Gaps 9;

OY 20 PATLPOQCAN-----KTCPTNYMNNHICRCLAQEDFMFSSDAGDSDTG 62
DB 1210 PPAFTCSASQRYSNVTCSCGCPGMPAPAGCGNOWNVCSQCCKPKMKRPAAN---- 1265
OY 63 TPGFHDICGPNKLEDETCQVCVRAGLRPASCGRPKELDRNSCCVC-KNKLFPSCGAN 121
DB 1266 -----CG-NKWNMDKACECECKRCPGACGKYQKMKNTCACECPGKATPPASCDDK 1317
OY 122 REDENTCQVCCKRT-----CPNQPPLNPGKACCTESPQCLLKGKFNHQTSCYR 176
DB 1318 KSNWPDSCSCQCKSKMPPGGCPSNQQWNCETCKCECS-GTQRC-PAGQSDSQTCQ-SC 1374
OY 177 PCTNRQKACRPGFSYSEYRCV 199
DB 1375 PATGR---CTGAQFWCAKCKCV 1394

RESULT 6
P97946 PRELIMINARY; PRT; 358 AA.
AC P97946;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR D (C-FOS, INDUCED GROWTH FACTOR).
GN VEGF-D OR FlgF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J;
RA MEDLINE=97030254; PubMed=8876195;
RA Orlandini M., Marconcini L., Ferruzzi R., Oliviero S.;
RT "Identification of a c-fos-induced gene that is related to the
RT platelet-derived growth factor/vascular endothelial growth factor
RT family.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:11675-11675(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RX MEDLINE=97349118; PubMed=9205122;
RA Yamada Y., Nezu J., Shimane M., Hirata Y.;
RT "Molecular cloning of a novel vascular endothelial growth factor,
RT VEGF-D.";
RL Genomics 42:483-488(1997).
DR EMBL; X99572; CAA67892.1; -.
DR EMBL; D89628; BA014002.1; -.
DR HSSP; P15692; IVP.
DR MGD; MGI:108037; FlgF.
DR InterPro; IPR000072; -.
DR Pfam; PF00341; PDGF_1.
DR ProDom; PD001629; -.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS0278; PDGF_2; 1.
DR SMART; SM00141; PDGF; 1.
SQ SEQUENCE 358 AA; 40908 MW; 6636B17FBF07037C CRC64;

Query Match 19.8%; Score 241.5; DB 11; Length 358;
Best Local Similarity 29.4%; Pred. No. 2e-18;
Matches 57; Conservative 23; Mismatches 57; Indels 57; Gaps 7;

OY 9 ROVHSIIRSLPA-TLPQCAANKTCPTNYMNNHICRCLAQEDFMFSSDAGDSDTG 67
DB 202 RHPYSIIRSLTQTPBEDCPHKKLCPIDMLDNTKCKCVLQDETPPL----- 248
OY 68 DICGPNKLEDETCQVCVRAGLRPASCGRPKELDRNSCCVCCKNKLFPSCGANREFD 127

Db 249 ----POTE-DHSLYIQ-----EPTLCGPH-----MTFDED 272
 QY 128 TCOVCYKRTCPRNQPLNPGKAC-ECTESPQKLLGKKFHHQTCSC-----YRRPCPN 180
 Db 273 RCEVCYKACPCGDLIOHPNCSCEFECKSLSCCKKHKLFHPDTCSEDRCPFHRTCAS 332
 QY 181 ROKACEPGFSYSEE 194
 Db 333 RKPACGKHWRFPE 346

RESULT 7

Q94438 PRELIMINARY; PRT; 1698 AA.
 AC Q94438;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE 185 KDA SILK PROTEIN.
 GN SPI85.
 OS Chironomus pallidivittatus (Midge).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
 OC Chironomidae; Chironomidae; Chironominae; Chironomus.
 OX NCBI_TaxID=7151;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary Gland;
 RA Case S.T., Cox C., Bell W.C., Hoffman R.T., Martin J., Hamilton R.;
 RU Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; U54640; AAA9803.1; -;
 DR InterPro; IPR000561; -;
 DR InterPro; IPR000853; -;
 DR PRINTS; PR00876; MTNEMATODE.
 DR PROSITE; PS00022; EGF_1; UNKNOMN_1.
 SQ SEQUENCE 1698 AA; 186164 MW; 82AF0838D9E2360E CRC64;

Query Match 19.3%; Score 235.5; DB 5; Length 1698;
 Best Local Similarity 25.7%; Pred. No. 3.3e-17;
 Matches 63; Conservative 32; Mismatches 79; Indels 71; Gaps 12;

QY 18 SLPALP-----QC--QAAN-KTCPTNVMNNHICLAQDFPMSSDAGDDTDGEF 66
 Db 1027 AMPAIPPEKKNEDKVCLECANVTCCEPQPCDNOCKCICPO-----V 1071
 QY 67 HDICGPNKLEDETCOC-----VCRAGLR-----PACGPHKELDRNSCQVC 109
 Db 1072 NPKCSDKQKFTSCCEGCDTQCKNGFRMSNLECGCLDEKKCKQKQVFPKNTQCKC 1131
 QY 110 KKKLPPOGAGNREFDETCOCVCK-----RTCPNQPILNPGKACAC-TESPQKLLG 163
 Db 1132 PQKRGDPCGNGKDCPLDSCCKSPKPRANGCPGVQEMNEKCCGCECPKPKKQCGG 1191
 QY 164 KKFHHQTCSC---YRRPCTNRQK-----ACEPGFS-----YSEEVCKCV-PS 201
 Db 1192 QDMNHLQCGGCPAPATPCSNKQKSNVSCSGCKPKPKRKGPKQIMCENTCRVCVPK 1251
 QY 202 YWKR 206
 Db 1252 NMCKP 1256

RESULT 8

Q35251 PRELIMINARY; PRT; 326 AA.
 AC Q35251;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR D.
 GN VEGF-D.
 Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE DAWLEY;
 RX MEDLINE=97349118; PubMed=9205122;
 RA Yamada Y., Nezu J., Shimane M., Hirata Y.;
 RT "Molecular cloning of a novel vascular endothelial growth factor,
 RT VEGF-D";
 RL Genomics 42:483-488(1997).
 DR EMBL; AF014827; AAB6557.1; -;
 DR HSSP; P15692; 1VP.
 DR InterPro; IPR000072; -;
 DR Pfam; PF00341; PDGF; 1.
 DR ProDom; PD001629; -; 1.
 DR PROSITE; PS00249; PDGF_1; 1.
 DR PROSITE; PS00278; PDGF_2; 1.
 DR SMART; SM00141; PDGF; 1.
 SQ SEQUENCE 326 AA; 37112 MW; 1261AFA373596C00 CRC64;

Query Match 17.9%; Score 217.5; DB 11; Length 326;
 Best Local Similarity 31.1%; Pred. No. 7.3e-16;
 Matches 52; Conservative 19; Mismatches 41; Indels 55; Gaps 7;

QY 9 ROVHSIIRSLPALP-----OCOANKTCPTNVMNNHICRCLAQDFMSSDAGDDTDG 65
 Db 202 RHPIYIIRSR--QIPREDQCHSKKLCIPVIMNDNTCKCYLD-----ENPLPG 250
 QY 66 PHDICGPNKLEDETCOCVCRAGLRPASCGRPHKELDRNSCQVCNKLPSQGANREFD 125
 Db 251 TED-----HSYLQE-----PALCGPH-----MMFD 270
 QY 126 EHTCGQVCKRCPRNQPLNPGKAC-ECTESPQKLLGKKFHHQTC 171
 Db 271 EDRCEVCYKACPCGDLIOHPNCSCEFECKSLSCCKKHKMFHPDTC 317

RESULT 9

Q74567 PRELIMINARY; PRT; 704 AA.
 AC Q74567;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE Q174 PROTEIN.
 GN Q174.
 OS Trichoderma harzianum.
 OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Trichoderma.
 OX NCBI_TaxID=5544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CECT 2413;
 RX MEDLINE=98263335; PubMed=9600944;
 RA Rey M., Ohno S.A., Pinior-Toro J.A., Jose A., Lobell A., Benitez T.;
 RT "Unexpected homology between inducible cell wall protein Q174 of
 RT filamentous fungi and BR3 salivary protein of the insect Chironomus";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:6212-6216(1998).
 DR EMBL; X95671; CA64974.1; -;
 DR InterPro; IPR000561; -;
 DR PROSITE; PS01186; EGF_2; 1.
 SQ SEQUENCE 704 AA; 77925 MW; 6341ABDECC365EBC CRC64;

Query Match 14.1%; Score 171.5; DB 3; Length 704;
 Best Local Similarity 22.8%; Pred. No. 1.4e-10;
 Matches 58; Conservative 26; Mismatches 67; Indels 103; Gaps 15;

QY 33 CPITNMMNNHICRCLAQDFMSSDAGDDTDGPHD-ICGPNKEL-----DEETCCV----- 84
 Db 78 CPSTYMWNGHOC-----VHDCGKDATWQYCNVCNKKRGEVYNPDKDTSCCPGQY 127

```
QY 85 -----CRAGLRPASCGRPHKELDRNSCCVCK--NKLFP-----Q 117
      |||
      |||
      |||
Db 128 WNGKNCQ-----VDCGKSDASYDKQKCKVCKHGEIYNSKTCSCPPQGVWNGVACYVD 182
QY 118 CCANREFEDNTCCVCKR-----TCPRNPPLPGKAC-----ECTE-- 154
      |||
      |||
      |||
Db 183 CCKEAFHFDKORCKCANNYNGEYNSGSKTACPGGQYFNGKKCVCPYGRVWNGKCCVEDC 242
QY 155 -----SPKCLLK--GKKFH--HOTSC-----YRRPCTNRQ-----KACEPG-- 188
      |||
      |||
      |||
Db 243 GKEAFHEDYKCKCCKVCKNNGEVYNSAKKTCSCPDGQYWNGKCCVCPYGVFNGKCCVPCDG 302
QY 189 ----FSYSEECVRC 198
      |||
      |||
      |||
Db 303 KEATFDYKCKCVC 316

RESULT 10
Q9GLX1 PRELIMINARY; PRT; 122 AA.
AC 09GLX1:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR-D (FRAGMENT).
OS VEGF-D.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG.
RA Manriota S.J., Pepper M.S.;
RA Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF099135; AAG29747.1;
FT NON_TER 1
FT NON_TER 122
FT NON_TER 122
SQ SEQUENCE 122 AA; 13820 MW; CC504B00E29D54EB CRC64;

Query Match
Best Local Similarity 12.6%; Score 153; DB 6; Length 122;
Matches 37; Conservative 24; Mismatches 28; Indels 56; Gaps 6;

QY 9 ROVHSIIRSLPATLPQ---COANKTCPTNMMNNHICRLAODEFMFSSDAGDSDTG 65
      |||
      |||
      |||
Db 30 RHPFSIIRSI--QIPEDRCSHKKRLCPVDMWNSKCKVLOEE----- 73
QY 66 FHDICGPNKLEDEETCCVCRAGLRPASCGRPH-KELDRNSCCVCKNKLFPSCGANREF 124
      |||
      |||
      |||
Db 74 -----NPLAGMDHT-----HLQEL-----ALCGQHNKF 97
QY 125 DENTCCVCKRTCPRNOPLPKCA 149
      |||
      |||
      |||
Db 98 DEDRCEVCCKTCCPRDLIOHPENC 122

RESULT 11
Q90824 PRELIMINARY; PRT; 1810 AA.
AC 090824:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE CYTOFACTIN 200 KDA PRECURSOR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=89184536; PubMed=2467292;
RA Jones F.S., Hoffman S., Cunningham B.A., Edelman G.M.;
RT "A detailed structural model of cytotactin: protein homologues,
RT alternative RNA splicing, and binding regions."
RL Proc. Natl. Acad. Sci. U.S.A. 86:1905-1909(1989).
DR EMBL; J04519; AAA48745.1;
DR HSSP; P24821; ITEN.
DR InterPro; IPR000561;
DR InterPro; IPR001777;
DR InterPro; IPR002181;
DR Pfam; PF00008; EGF_13.
DR Pfam; PF00041; fn3; 11.
DR Pfam; PF00147; fibrinogen C; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_14.
DR PROSITE; PS01186; EGF_2; 14.
DR SMART; SM00186; FBG; 1.
KW EGF-like domain; Glycoprotein; Signal.
FT SIGNAL 1
FT CHAIN 33 1810
SQ SEQUENCE 1810 AA; 199304 MW; 42B5FEC7BA181DAC CRC64;

Query Match
Best Local Similarity 12.4%; Score 151; DB 13; Length 1810;
Matches 61; Conservative 21; Mismatches 78; Indels 92; Gaps 18;

QY 24 PCCQANKTCPTNMMNNHIC---RCLAODEF-----MFSSDAGDS--TDG---F 66
      |||
      |||
      |||
Db 185 PNC--SEPCAPRNCI-NRGLCVRAKICEEGFTGEDCSQARPCDNCQKCVDCVCF 241
QY 67 HDICGPNKLEDEETC-----QCVRAGLRPASC-----PHKELDRNSC---Q 106
      |||
      |||
      |||
Db 242 EGYTGP--DCGELCPHGGCIHRCVCGRCVCHGFTGEDCNEPLCPNCHRGKCVDNE 299
QY 107 CVCKNKLFPSCG---ANREFDNTC---QCVR-----TCPRNPPLN---P 145
      |||
      |||
      |||
Db 300 CVCDGYTGEDCGELICPNDCFDRGRCINGTCFCEGYTGEDCGELTCNNNGNGRCN 359
QY 146 GKAC-----ECTES--PQKCLKKKFHHQTSCT-----RRP 177
      |||
      |||
      |||
Db 360 GLCVCHGEGFVGDSCQKRCPTCNRRVCVDGR-CVCHGTYLGEDCGELICPNDCHNRGR 418
QY 178 CTRNQKACEPGF 189
      |||
      |||
      |||
Db 419 CINGCCVDCDEGF 430

RESULT 12
Q9VB78 PRELIMINARY; PRT; 800 AA.
AC 09VB78:
DT 01-MAR-2000 (TREMBLrel. 13, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE CG6124 PROTEIN.
CN CG6124.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton R.G., Wotman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
```


[2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=90030407; PubMed=2478295;
 RX Spring J., Beck K.F., Chiquet-Ehrismann R.;
 RT "Two contrary functions of tenascin: dissection of the active sites by
 RL recombinant tenascin fragments.";
 RL Cell 59:325-334(1989).
 DR EMBL: M23121; AAA9084.1; -
 DR HSSP: P24821; ITEN
 DR InterPro: IPR000561; -
 DR InterPro: IPR001777; -
 DR InterPro: IPR002181; -
 DR Pfam: PF00008; EGF_13.
 DR Pfam: PF00041; fn3_10.
 DR Pfam: PF00147; fibrinogen_C_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_14.
 DR PROSITE: PS01186; EGF_2; 14.
 DR SMART: SM01186; FBG; 1.
 DR EGF-like domain; glycoprotein; signal.
 FT SIGNAL 1 33 POTENTIAL.
 FT CHAIN 34 1532 POTENTIAL.
 FT CHAIN 34 1224 POTENTIAL.
 FT CHAIN 34 1714 POTENTIAL.
 FT CHAIN 34 1042 POTENTIAL.
 FT CHAIN 1225 1714 POTENTIAL.
 SQ SEQUENCE 1714 AA; 188515 MW; A6FB0CFD623CDE53 CRC64;

Query Match 12.2%; Score 148; DB 13; Length 1714;
 Best Local Similarity 24.2%; Pred. No. 1.1e-07;
 Matches 61; Conservative 20; Mismatches 79; Indels 92; Gaps 18;

QY 24 PCCQAANKTCPTNYMNNHIC---RCLAOEDF-----MESSAGDS---TDG-----F 66
 Db 165 PNC--SEAPCPNCL--NGLCYRKCICEBEGFTGEDCSOACPSDCNQGKVCYDVCVF 241
 QY 67 HDICSPNKELEDETC-----QCYCRAGLRPASCG---PKKELDRNSC---Q 106
 Db 242 EGYTGP--DCGELCPHCGIHGRVCVGRVCYCEGFTGEDCNELPLCPNCHNRGCYVNE 299
 QY 107 CYCKRKLFPSSCG---ANREPDENTC---QCVCNR-----TCPRNPLN---P 145
 Db 300 CVCDEGYTGEDGELICPNDECFGRGRCINGTCFCEGYTGEDGELTCTPNNGNGRCEN 359
 QY 146 GKAC-----ECTES--POKCLGKKKFFHOTGSCY-----RRP 177
 Db 360 GLCYCHHEGFVGDGCKQKRCPRKCNRRGHCYDGR-CVCHEGYLGEDGELRCPNCHNRGR 418
 QY 178 CTNRKACEPGF 189
 Db 419 CINGCVCDEGF 430

RESULT 15
 Q9VJ05 PRELIMINARY; PRT; 543 AA.
 AC Q9VJ05;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE BG:DS00180.10 PROTEIN.
 GN BG:DS00180.10 OR CG8944.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI TaxID=7227;
 OX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amaratilakes P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Ballwey R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Broksstein P., Brottier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Genter A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheller F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003642; AAF53364.1; -
 DR HSSP: P35555; IEMN.
 DR FlyBase: FBgn0028545; BG:DS00180.10.
 DR InterPro: IPR000561; -
 DR InterPro: IPR001304; -
 DR PROSITE: PS00615; C-TYPE LECTIN_1; UNKNOWN_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS01186; EGF_2; 7.
 DR SMART: SM00181; EGF_1.
 KM EGF-like domain; glycoprotein.
 SQ SEQUENCE 543 AA; 58562 MW; F419FDAFC4AD0D2A CRC64;

Query Match 11.9%; Score 145; DB 5; Length 543;
 Best Local Similarity 23.4%; Pred. No. 8.7e-08;
 Matches 57; Conservative 19; Mismatches 98; Indels 70; Gaps 15;

QY 18 SLPATLPCCOA-----ANKTCPTNYMNNHIC---CLAODEFMSSDAGD-----DS 62
 Db 31 SSPNTC-SCNAGYGGIDCHPVCPT-VCGKNERCDBRGVCSQNGYKXRPSPDNCLPVCEK 88
 QY 63 TDFHDTGPNKELEDETCQCYCRAGLRPASCG---HKELDNRSCQCYCKNKLFPSSQ 117
 Db 89 ECGHHSFC-----SEPGKCECEPGEYKVGNGTFFPDGYKNNNGNCSPLC-----PKD 136
 QY 118 GGAN-REPDENTCOC-----VCKRCPRN-QPLNGKACE-----151
 Db 137 CGONSRCVPRGVCCEGNCYAGDDGCTNCRPVCS-TCPENGICLSPGVCVCRGYVMRNDL 195
 QY 152 CTESPOKCLGKKRFFHOTGSCY-----RRPCTNR-QKACEPGFSYSEVCRCPVSY 202
 Db 196 CQPHCEKCSDNAHCYAPNQCCEPFGYESSGADKKCYPRCSKGCINGRFAFETCVCSIGY 255
 QY 203 WKRP 206
 Db 256 QMGF 259

Wed Oct 17 15:32:22 2001

us-09-427-657-2_copy_211_419.rspt

Page 8

Search completed: October 17, 2001, 14:50:53
Job time: 419 sec

0 2. 2.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:46:00 ; Search time 115.93 seconds
(without alignments)
100.404 Million cell updates/sec

Title: US-09-427-657-2_COPY_228_419

Perfect score: 1130
Sequence: 1 SLPLATLPGCAANKTCPTNY.....SYSEVRCRCVPSYWKRPQMS 192

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
A.Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1130	100.0	292	20	AAW86225
2	1130	100.0	297	20	AAW86224
3	1130	100.0	302	20	AAW86223
4	1130	100.0	307	20	AAW86222
5	1130	100.0	318	20	AAW86221
6	1130	100.0	329	20	AAW86220
7	1130	100.0	339	20	AAW86219
8	1130	100.0	349	18	AAW86218
9	1130	100.0	359	18	AAW86217
10	1130	100.0	369	18	AAW86216
11	1130	100.0	379	19	AAW86215

12	1130	100.0	419	20	AAW86203	Human vascular end
13	1130	100.0	419	21	AAW86204	Human VEGF-C prote
14	1130	100.0	419	21	AAW86205	Human VEGF-C prote
15	1130	100.0	419	21	AAW86206	Human VEGF-C prote
16	1130	100.0	419	21	AAW86207	Human VEGF-C prote
17	1130	100.0	419	22	AAW86208	Human VEGF-C prote
18	1130	100.0	419	22	AAW86209	Human VEGF-C prote
19	1130	100.0	419	22	AAW86210	Human VEGF-C prote
20	1130	100.0	419	22	AAW86211	Human VEGF-C prote
21	1130	100.0	419	22	AAW86212	Human VEGF-C prote
22	1130	100.0	419	22	AAW86213	Human VEGF-C prote
23	1130	100.0	419	22	AAW86214	Human VEGF-C prote
24	1130	100.0	419	22	AAW86215	Human VEGF-C prote
25	1130	100.0	419	22	AAW86216	Human VEGF-C prote
26	1130	100.0	419	22	AAW86217	Human VEGF-C prote
27	1130	100.0	419	22	AAW86218	Human VEGF-C prote
28	1130	100.0	419	22	AAW86219	Human VEGF-C prote
29	1130	100.0	419	22	AAW86220	Human VEGF-C prote
30	1130	100.0	419	22	AAW86221	Human VEGF-C prote
31	1130	100.0	419	22	AAW86222	Human VEGF-C prote
32	1130	100.0	419	22	AAW86223	Human VEGF-C prote
33	1130	100.0	419	22	AAW86224	Human VEGF-C prote
34	1130	100.0	419	22	AAW86225	Human VEGF-C prote
35	1130	100.0	419	22	AAW86226	Human VEGF-C prote
36	1130	100.0	419	22	AAW86227	Human VEGF-C prote
37	1130	100.0	419	22	AAW86228	Human VEGF-C prote
38	1130	100.0	419	22	AAW86229	Human VEGF-C prote
39	1130	100.0	419	22	AAW86230	Human VEGF-C prote
40	1130	100.0	419	22	AAW86231	Human VEGF-C prote
41	1130	100.0	419	22	AAW86232	Human VEGF-C prote
42	1130	100.0	419	22	AAW86233	Human VEGF-C prote
43	1130	100.0	419	22	AAW86234	Human VEGF-C prote
44	1130	100.0	419	22	AAW86235	Human VEGF-C prote
45	1130	100.0	419	22	AAW86236	Human VEGF-C prote

ALIGNMENTS

RESULT 1	AAW86225	standard; protein; 292 AA.
ID	AAW86225	
XX	AAW86225	
AC	AAW86225	
XX	AAW86225	
DT	16-FEB-1999	(first entry)
DE	Human VEGF-C truncated fragment 4.	
XX	VEGF: VEGF: vascular endothelial growth factor; VEGF-related protein;	
KW	recombinant; truncated; gene therapy; angiogenesis; cardiac ischemia;	
KW	coronary; collateral vessel development; cell growth; migration; heart;	
KW	lower limb ischemia; stroke; peripheral vascular disease; intestine;	
KW	wound healing; skin; vascular permeability.	
OS	Homo sapiens.	
XX	W09849300-A2.	
XX	W09849300-A2.	
PD	05-NOV-1998.	
XX	20-APR-1998;	98WO-US07801.
XX	25-APR-1997;	97US-0842984.
PA	(COLL-) COLLATERAL THERAPEUTICS.	
XX	Bohlen P;	
XX	WPI; 1999-009426/01.	
DR	New truncated vascular endothelial growth factor-related protein	
XX	subunits - lack part of the N-terminal sequence, used to stimulate	
PT		

PT angiogenesis, e.g. for treating heart disease and ischaemia
XX
PS Claim 5; Fig 2D; 113pp; English.
XX
CC The invention relates to truncated VRP (vascular endothelial growth
CC factor (VEGF)-related protein) subunits that have at least one amino
CC acid N-terminal to the first Cys of the core sequence deleted. Host
CC cells transformed or transfected with expression vectors containing
CC nucleic acids encoding the truncated VRP subunits are used to produce
CC the truncated proteins recombinantly. The truncated VRP subunits,
CC optionally expressed from gene therapy vectors, have in vivo and in vitro
CC angiogenic activity and are used to stimulate angiogenesis, particularly
CC coronary collateral vessel development in cases of cardiac ischaemia; to
CC stimulate endothelial cell growth and migration in vitro; to treat heart
CC disease; to treat ischaemia (e.g. cardiac, chronic coronary or chronic
CC lower limb ischaemia; stroke and peripheral vascular disease); to promote
CC healing of wounds (of skin or intestines), and to increase vascular
CC permeability. Sequences AAW86222 to AAW86225 represent truncated
CC fragments of VEGF-C.
XX
SQ Sequence 292 AA:

Query Match 100.0%; Score 1130; DB 20; Length 292;
Best Local Similarity 100.0%; Pred. No. 2.2e-82;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLPTLPQCOAANKTCPTNYVMNNHICRCLAOEDFMFSSDADGSDTDFHDCGPKKELD 60
DB 101 slptlpqcaanktcptnyvmnnhircrlagedtmfssdagdsdtdghdicspnkeld 160
OY 61 EETCQCVCRAGLRPASCGPHKELDRNSCQVCYCKNKLFPSCGANKREPDENTCQVCYCKRTC 120
DB 161 eetcqcvcraglrpsacgphkeldrnsqcvcknklfpscgankreidentcqvckrtc 220
OY 121 PRNQPPLNPGKACCECTESQKCLLGKKFHHOTGCTYRRPCTNRQKACBPGRSYSEVCR 180
DB 221 prngplmpgkacectespqckllgkktfhqtcscyrpctnrqkacepgrfsysevcr 280
OY 181 CVPSTYMKRPOMS 192
DB 281 cvpsywkrpoms 292

RESULT 2
AAW86224
ID AAW86224 standard; protein; 297 AA.
XX
AC AAW86224;
XX
DT 16-FEB-1999 (first entry)
XX
DE Human VEGF-C truncated fragment 3.
XX
KW VEGF: VRP: vascular endothelial growth factor; VEGF-related protein;
KW recombinant; truncated; gene therapy; angiogenesis; cardiac ischaemia;
KW coronary; collateral vessel development; cell growth; migration; heart;
KW lower limb ischaemia; stroke; peripheral vascular disease; intestine;
KW wound healing; skin; vascular permeability.
XX
OS Homo sapiens.
XX
PN WO9849300-A2.
XX
PD 05-NOV-1998.
XX
PE 20-APR-1998; 98WO-US07801.
XX
PR 25-APR-1997; 97US-0842984.
XX
PA (COLL-) COLLATERAL THERAPEUTICS.
XX
PI Bohlen P;

XX
DR WPI: 1999-009426/01.
XX
PT New truncated vascular endothelial growth factor-related protein
PT subunits - lack part of the N-terminal sequence, used to stimulate
PT angiogenesis, e.g. for treating heart disease and ischaemia
XX
PS Claim 5; Fig 2D; 113pp; English.
XX
CC The invention relates to truncated VRP (vascular endothelial growth
CC factor (VEGF)-related protein) subunits that have at least one amino
CC acid N-terminal to the first Cys of the core sequence deleted. Host
CC cells transformed or transfected with expression vectors containing
CC nucleic acids encoding the truncated VRP subunits are used to produce
CC the truncated proteins recombinantly. The truncated VRP subunits,
CC optionally expressed from gene therapy vectors, have in vivo and in vitro
CC angiogenic activity and are used to stimulate angiogenesis, particularly
CC coronary collateral vessel development in cases of cardiac ischaemia; to
CC stimulate endothelial cell growth and migration in vitro; to treat heart
CC disease; to treat ischaemia (e.g. cardiac, chronic coronary or chronic
CC lower limb ischaemia; stroke and peripheral vascular disease); to promote
CC healing of wounds (of skin or intestines), and to increase vascular
CC permeability. Sequences AAW86222 to AAW86225 represent truncated
CC fragments of VEGF-C.
XX
SQ Sequence 297 AA:

Query Match 100.0%; Score 1130; DB 20; Length 297;
Best Local Similarity 100.0%; Pred. No. 2.2e-82;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLPTLPQCOAANKTCPTNYVMNNHICRCLAOEDFMFSSDADGSDTDFHDCGPKKELD 60
DB 106 slptlpqcaanktcptnyvmnnhircrlagedtmfssdagdsdtdghdicspnkeld 165
OY 61 EETCQCVCRAGLRPASCGPHKELDRNSCQVCYCKNKLFPSCGANKREPDENTCQVCYCKRTC 120
DB 166 eetcqcvcraglrpsacgphkeldrnsqcvcknklfpscgankreidentcqvckrtc 225
OY 121 PRNQPPLNPGKACCECTESQKCLLGKKFHHOTGCTYRRPCTNRQKACBPGRSYSEVCR 180
DB 226 prngplmpgkacectespqckllgkktfhqtcscyrpctnrqkacepgrfsysevcr 285
OY 181 CVPSTYMKRPOMS 192
DB 286 cvpsywkrpoms 297

RESULT 3
AAW86223
ID AAW86223 standard; protein; 302 AA.
XX
AC AAW86223;
XX
DT 16-FEB-1999 (first entry)
XX
DE Human VEGF-C truncated fragment 2.
XX
KW VEGF: VRP: vascular endothelial growth factor; VEGF-related protein;
KW recombinant; truncated; gene therapy; angiogenesis; cardiac ischaemia;
KW coronary; collateral vessel development; cell growth; migration; heart;
KW lower limb ischaemia; stroke; peripheral vascular disease; intestine;
KW wound healing; skin; vascular permeability.
XX
OS Homo sapiens.
XX
PN WO9849300-A2.
XX
PD 05-NOV-1998.
XX
PE 20-APR-1998; 98WO-US07801.
XX

PR 25-APR-1997; 97US-0842984.
XX
XX (COLL-) COLLATERAL THERAPEUTICS.
XX
XX Bohlen P;
PI
XX WPI; 1999-009426/01.
XX
XX New truncated vascular endothelial growth factor-related protein
PT subunits - lack part of the N-terminal sequence, used to stimulate
PT angiogenesis, e.g. for treating heart disease and Ischaemia
XX
XX
PS Claim 5; Fig 2D; 113pp; English.
XX
XX The invention relates to truncated VRP (vascular endothelial growth
CC factor (VEGF)-related protein) subunits that have at least one amino
CC acid N-terminal to the first Cys of the core sequence deleted. Host
CC cells transformed or transfected with expression vectors containing
CC nucleic acids encoding the truncated VRP subunits are used to produce
CC the truncated proteins recombinantly. The truncated VRP subunits,
CC optionally expressed from gene therapy vectors, have in vivo and in vitro
CC angiogenic activity and are used to stimulate angiogenesis, particularly
CC coronary collateral vessel development in cases of cardiac Ischaemia; to
CC stimulate endothelial cell growth and migration in vitro; to treat heart
CC disease; to treat Ischaemia (e.g. cardiac, chronic coronary or chronic
CC lower limb Ischaemia; stroke and peripheral vascular disease); to promote
CC healing of wounds (of skin or intestines), and to increase vascular
CC permeability. Sequences AAM86222 to AAM86225 represent truncated
CC fragments of VEGF-C.
XX
XX Sequence 302 AA:
SQ

Query Match 100.0%; Score 1130; DB 20; Length 302;
Best Local Similarity 100.0%; Pred. No. 2.3e-82;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPATLPQCAANKTCPTVYMNHNHICRCLAOEDFEFSSDAGDSDTGDHICGPNKEID 60
DB 111 s|p|a|t|p|q|g|a|n|k|t|c|p|t|v|y|m|n|h|n|h|i|c|r|c|l|a|o|e|d|f|e|f|s|s|d|a|g|d|s|d|t|g|d|h|i|c|g|p|n|k|e|i|d| 170
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 EETCCVCVRAGLRPASCGPHKELDNRSCOCVCKNKLFPSSQGANREFDENTCQVCCKRTC 120
DB 171 e|e|t|c|c|v|c|r|a|g|l|r|p|a|s|c|g|p|h|k|e|l|d|n|r|s|c|o|c|v|c|k|n|k|l|f|p|s|s|q|a|n|r|e|f|e|n|t|c|q|v|c|k|r|t|c| 230
QY 121 PRNQPINPEKCKACECTESQKCLKGKKRHHOTCSYRRPCTNRQKACPRGFSYSEVCR 180
DB 231 p|r|n|q|p|i|n|p|e|k|c|a|c|e|t|e|s|q|k|c|l|k|g|k|k|r|h|h|o|t|c|s|y|r|r|p|c|t|n|r|q|k|a|c|p|r|g|f|s|y|s|e|v|c|r| 290
QY 181 CVPSYWKRPOMS 192
DB 291 c|v|p|s|y|w|k|r|p|o|m|s| 302

RESULT 4
AAM86222
ID AAM86222 standard; protein: 307 AA.
XX
XX AAM86222;
AC
XX 16-FEB-1999 (first entry)
DT
XX
XX Human VEGF-C truncated fragment 1.
DE
XX
XX VEGF; VRP; vascular endothelial growth factor; VEGF-related protein;
KW recombinant; truncated; gene therapy; angiogenesis; cardiac Ischaemia;
KW coronary; collateral vessel development; cell growth; migration; heart;
KW lower limb Ischaemia; stroke; peripheral vascular disease; intestine;
KW wound healing; skin; vascular permeability.
XX
XX Homo sapiens.
OS
XX
XX WO9849300-A2.
PN

XX
XX 05-NOV-1998.
PD
XX
XX 20-APR-1998; 98WO-US07801.
PF
XX
XX 25-APR-1997; 97US-0842984.
PR
XX
XX (COLL-) COLLATERAL THERAPEUTICS.
PA
XX Bohlen P;
PI
XX WPI; 1999-009426/01.
XX
XX New truncated vascular endothelial growth factor-related protein
PT subunits - lack part of the N-terminal sequence, used to stimulate
PT angiogenesis, e.g. for treating heart disease and Ischaemia
XX
XX
PS Claim 5; Fig 2D; 113pp; English.
XX
XX The invention relates to truncated VRP (vascular endothelial growth
CC factor (VEGF)-related protein) subunits that have at least one amino
CC acid N-terminal to the first Cys of the core sequence deleted. Host
CC cells transformed or transfected with expression vectors containing
CC nucleic acids encoding the truncated VRP subunits are used to produce
CC the truncated proteins recombinantly. The truncated VRP subunits,
CC optionally expressed from gene therapy vectors, have in vivo and in vitro
CC angiogenic activity and are used to stimulate angiogenesis, particularly
CC coronary collateral vessel development in cases of cardiac Ischaemia; to
CC stimulate endothelial cell growth and migration in vitro; to treat heart
CC disease; to treat Ischaemia (e.g. cardiac, chronic coronary or chronic
CC lower limb Ischaemia; stroke and peripheral vascular disease); to promote
CC healing of wounds (of skin or intestines), and to increase vascular
CC permeability. Sequences AAM86222 to AAM86225 represent truncated
CC fragments of VEGF-C.
XX
XX Sequence 307 AA:
SQ

Query Match 100.0%; Score 1130; DB 20; Length 307;
Best Local Similarity 100.0%; Pred. No. 2.3e-82;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPATLPQCAANKTCPTVYMNHNHICRCLAOEDFEFSSDAGDSDTGDHICGPNKEID 60
DB 116 s|p|a|t|p|q|g|a|n|k|t|c|p|t|v|y|m|n|h|n|h|i|c|r|c|l|a|o|e|d|f|e|f|s|s|d|a|g|d|s|d|t|g|d|h|i|c|g|p|n|k|e|i|d| 175
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 EETCCVCVRAGLRPASCGPHKELDNRSCOCVCKNKLFPSSQGANREFDENTCQVCCKRTC 120
DB 176 e|e|t|c|c|v|c|r|a|g|l|r|p|a|s|c|g|p|h|k|e|l|d|n|r|s|c|o|c|v|c|k|n|k|l|f|p|s|s|q|a|n|r|e|f|e|n|t|c|q|v|c|k|r|t|c| 235
QY 121 PRNQPINPEKCKACECTESQKCLKGKKRHHOTCSYRRPCTNRQKACPRGFSYSEVCR 180
DB 236 p|r|n|q|p|i|n|p|e|k|c|a|c|e|t|e|s|q|k|c|l|k|g|k|k|r|h|h|o|t|c|s|y|r|r|p|c|t|n|r|q|k|a|c|p|r|g|f|s|y|s|e|v|c|r| 295
QY 181 CVPSYWKRPOMS 192
DB 296 c|v|p|s|y|w|k|r|p|o|m|s| 307

RESULT 5
AAY08284
ID AAY08284 standard; Protein: 318 AA.
XX
XX AAY08284;
AC
XX 14-JUL-1999 (first entry)
DT
XX
XX Human growth factor protein fragment VEGF-C.
DE
XX
XX Human growth factor; human; dimer; cysteine knot; cellular inclusion body;
KW pharmaceutical.
KW
XX
XX Homo sapiens.
OS

XX DE19748734-AL.
 PN
 XX
 PD 06-MAY-1999.
 XX
 PF 05-NOV-1997; 97DE-1048734.
 XX
 PR 05-NOV-1997; 97DE-1048734.
 XX
 PA (GBFB) GES BIOTECHNOLOGISCHE FORSCHUNG MBH.
 XX
 PI Erdmann H, Kaerst U, Mueller C, Rinas U, Weich H;
 XX
 DR WPI: 1999-278785/24.
 XX
 PF Preparing active growth factor dimers from inclusion bodies in high
 PT yield
 XX
 PS Claim 14; Page 8; 14pp; German.
 XX
 CC This invention describes the novel preparation of biologically active
 CC dimers of recombinant human growth factors of the cysteine knot family
 CC starting from cellular inclusion bodies. Such dimers are useful in
 CC pharmaceutical compositions and the method provides yields of 31-39.7%.
 CC In examples, compared with about 10% for the conventional method (see
 CC Biochemistry, 28 (1989) 2956). AAY08278-Y08301 are human growth factor
 CC protein fragments used in the method of the invention.
 XX
 SO Sequence 318 AA;

Query Match 100.0%; Score 1130; DB 20; Length 318;
 Best Local Similarity 100.0%; Pred. No. 2.4e-82;
 Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLPATLPQCOAANKTCPTNYMNNHICRCLAOEDFMFSSDADSDTDGPHDICGPNKEID 60
 |||||||
 DB 127 sLPATLPQGAANKTCPTNYMNNHICRCLAQEDFMFSSDADSDTDGPHDICGPNKEID 186
 QY 61 EETCCOVCVAGLRPAASCGPHKELDRNSQCYVCNKLFPSCGANEFPDENTCQCYCKRTC 120
 |||||||
 DB 187 eETCGVCVAGLRPAASCGPHKELDRNSQCYVCNKLFPSCGANEFPDENTCQCYCKRTC 246
 QY 121 PRNQLPNPGKACECTESPQKCLKGKKFHHQTCSCYRRPCTNRQKACEPGFSYSEEVCR 180
 |||||||
 DB 247 PRNGPLNPJGKACECTESPQKCLKGKKFHHQTCSCYRRPCTNRQKACEPGFSYSEEVCR 306
 QY 181 CVPSTWKRPPQMS 192
 |||||||
 DB 307 CVPSTWKRPPQMS 318

RESULT 6
 AAW86237
 ID AAW86237 standard; protein: 399 AA.
 XX
 AC AAW86237;
 XX
 DT 16-FEB-1999 (first entry)
 XX
 DE Human VEGF-C full length sequence.

XX VEGF; VRF; vascular endothelial growth factor; VEGF-related protein;
 KW recombinant; truncated; gene therapy; angiogenesis; cardiac ischaemia;
 KW coronary; collateral vessel development; cell growth; migration; heart;
 KW lower limb ischaemia; stroke; peripheral vascular disease; intestine;
 KW wound healing; skin; vascular permeability.
 XX
 OS Homo sapiens.
 XX
 PN W09849300-A2.
 XX
 PD 05-NOV-1998.

XX 20-APR-1998; 98WO-US07801.
 PF
 XX
 PR 25-APR-1997; 97US-0842984.
 XX
 PA (COLL-) COLLATERAL THERAPEUTICS.
 XX
 PI Bohlen P;
 XX
 DR WPI: 1999-009426/01.
 XX

PT New truncated vascular endothelial growth factor-related protein
 PT subunits - lack part of the N-terminal sequence, used to stimulate
 PT angiogenesis, e.g. for treating heart disease and ischaemia
 XX
 PS Claim 5; Fig 2D; 113pp; English.

XX The invention relates to truncated VRF (vascular endothelial growth
 CC factor (VEGF)-related protein) subunits that have at least one amino
 CC acid N-terminal to the first Cys of the core sequence deleted. Host
 CC cells transformed or transfected with expression vectors containing
 CC nucleic acids encoding the truncated VRF subunits are used to produce
 CC the truncated proteins recombinantly. The truncated VRF subunits,
 CC optionally expressed from gene therapy vectors, have in vivo and in vitro
 CC angiogenic activity and are used to stimulate angiogenesis, particularly
 CC coronary collateral vessel development in cases of cardiac ischaemia; to
 CC stimulate endothelial cell growth and migration in vitro; to treat heart
 CC disease; to treat ischaemia (e.g. cardiac, chronic coronary or chronic
 CC lower limb ischaemia, stroke and peripheral vascular disease); to promote
 CC healing of wounds (of skin or intestines), and to increase vascular
 CC permeability. Sequences AAW86234 to AAW86239 represent full length VRF
 CC sequences from which the truncated fragments are created.

SO Sequence 399 AA;
 Query Match 100.0%; Score 1130; DB 20; Length 399;
 Best Local Similarity 100.0%; Pred. No. 2.9e-82;
 Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLPATLPQCOAANKTCPTNYMNNHICRCLAOEDFMFSSDADSDTDGPHDICGPNKEID 60
 |||||||
 DB 208 sLPATLPQGAANKTCPTNYMNNHICRCLAQEDFMFSSDADSDTDGPHDICGPNKEID 267
 QY 61 EETCCOVCVAGLRPAASCGPHKELDRNSQCYVCNKLFPSCGANEFPDENTCQCYCKRTC 120
 |||||||
 DB 268 eETCGVCVAGLRPAASCGPHKELDRNSQCYVCNKLFPSCGANEFPDENTCQCYCKRTC 327
 QY 121 PRNQLPNPGKACECTESPQKCLKGKKFHHQTCSCYRRPCTNRQKACEPGFSYSEEVCR 180
 |||||||
 DB 328 PRNGPLNPJGKACECTESPQKCLKGKKFHHQTCSCYRRPCTNRQKACEPGFSYSEEVCR 387
 QY 181 CVPSTWKRPPQMS 192
 |||||||
 DB 388 CVPSTWKRPPQMS 399

RESULT 7
 AAW17837
 ID AAW17837 standard; Protein: 419 AA.
 XX
 AC AAW17837;
 XX
 DT 13-JAN-1998 (first entry)
 XX
 DE Human foetal liver kinase A binding protein flk-1bp.

XX Foetal liver kinase 1 binding protein; human; flk-1bp;
 KW receptor tyrosine kinase; vasculogenesis; angiogenesis;
 KW wound healing; tumour; therapy; antagonist; antibody.
 XX
 OS Homo sapiens.
 XX

Key Location/Qualifiers
FH 1..20
FT /Label= Sig_peptide
FT 21..419
FT Protein
FT /Label= Mat_protein
FT /note= "(Claim 10)"
FT Peptide
FT 21..35
FT /Label= N-terminal
FT /note= "(Claim 9)"
XX
XX MO9717442-A1.
XX
XX 15-MAY-1997.
XX
XX 05-NOV-1996; 96MO-US17584.
XX
XX 08-NOV-1995; 95US-0554374.
XX
XX (IMMV) IMMUNEX CORP.
XX
XX Lyman SD;
XX
XX WPI: 1997-281031/25.
XX
XX N-PSDB; AAT68811.
XX
XX DNA encoding a human foetal liver kinase 1 binding protein - used
XX to treat conditions with insufficient protein, deliver agents to
XX cells and identify antagonists to treat protein-mediated conditions
XX
XX Claim 1; Page 30-32; 43pp; English.
XX
XX This polypeptide comprises a human foetal liver kinase 1 binding
XX protein (flk-1bp) (see AAM17837) that binds to the receptor tyrosine
XX kinase flk-1 expressed on vascular endothelial and other cells.
XX The mature flk-1bp can be secreted from host cells transformed with
XX an expression vector including an isolated flk-1bp cDNA clone (see
XX AAT68811). flk-1bp can be used to isolate cells to which it binds,
XX for use in studying the roles of such cells and of flk-1 in
XX vasculogenesis and angiogenesis. Angiogenesis inhibition or
XX increased vasularisation may be clinically desirable (e.g. to
XX suppress solid tumour growth or in wound healing, respectively).
XX The flk-1bp can be administered to treat conditions with defective
XX or insufficient flk-1. Polypeptides may also act as carriers to
XX deliver diagnostic/therapeutic agents to cells to which flk-1bp
XX binds, to generate antibodies, and to identify flk-1bp antagonists
XX useful for treating flk-1bp mediated conditions.
XX
XX Sequence 419 AA:
SQ
Query Match 100.0%; Score 1130; DB 18; Length 419;
Best Local Similarity 100.0%; Pred. No. 3.1e-82;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLPLATLPOQAAANKTPTNYMMNNHICRLAOEDFMFSSDADSDTDGFHDICGPNKELD 60
DB 228 slpaltlpgqaaanktcpnymmnnhlcrlagedmfssdadsdtdgfhdicgpnkeld 287
QY 61 EETCGCVCAGCAGRPASGCBPHKELDRNSCCQVCVKNLFPSSQCANREPDNTCCVCCKRTC 120
DB 288 eetcgcvcagcagrpasgcbphkeldrnsccvcvknlfpsqcanrepdntccvcckrtc 347
QY 121 PANQPLNPGKACCECTESPOKCLLKGGKFFHQTGSCYRPPCNRKACGPGSYSEEVCR 180
DB 348 ptnqplnpgkacectespqkcllkgkffhqtgscyrppcnrkacpgsyseevcr 407
QY 181 CVPSTYKRPQMS 192
DB 408 cvpsywkripqms 419
RESULT 8
AAM00932

ID AAM00932 standard; Protein; 419 AA.
XX
XX AAM00932;
AC 10-NOV-1997 (first entry)
XX
XX 10-NOV-1997 (first entry)
DT
XX
XX Human Flt4 receptor tyrosine kinase ligand VEGF-C.
DE
XX
XX VEGF-C; Flt4; receptor tyrosine kinase; VEGFR-3; human;
KW vascular endothelial growth factor receptor-3; ligand;
KW angiogenesis; wound healing; lymph vessel; lymphangioma;
KW cancer; metastasis; therapy; diagnosis; antibody; inhibitor.
OS
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 1..102
FT /Label= Prepro_peptide
FT 32..227
FT Peptide
FT /note= "preferred active fragment of VEGF-C,
FT retaining Flt4 ligand activity (Claim 15)"
FT Peptide
FT /note= "preferred active fragment of VEGF-C,
FT retaining Flt4 ligand activity (Claim 12)"
FT Peptide
FT /note= "preferred active fragment of VEGF-C,
FT retaining Flt4 ligand activity (Claim 13)"
FT Peptide
FT /note= "preferred active fragment of VEGF-C,
FT retaining Flt4 ligand activity (Claim 14)"
FT Peptide
FT /note= "preferred active fragment of VEGF-C,
FT retaining Flt4 ligand activity (Claim 10)"
FT Peptide
FT /note= "preferred active fragment of VEGF-C,
FT retaining Flt4 ligand activity (Claim 11)"
FT Peptide
FT /note= "preferred active fragment of VEGF-C,
FT retaining Flt4 ligand activity (Claim 9)"
FT Peptide
FT /note= "preferred active fragment of VEGF-C,
FT retaining Flt4 ligand activity (Claim 8)"
XX
XX MO9705250-A2.
XX
XX 13-FEB-1997.
XX
XX 01-AUG-1996; 96MO-FI00427.
XX
XX 28-JUN-1996; 96US-0671573.
XX
XX 01-AUG-1995; 95US-0510133.
XX
XX 12-JAN-1996; 96US-0585895.
XX
XX 14-FEB-1996; 96US-0601132.
XX
XX (UYHE-) UNIV HELSINKI LICENSING LTD OY.
XX
XX Alltalo K, Joukov V;
XX
XX WPI: 1997-145688/13.
XX
XX N-PSDB; AAT84276.
XX
XX Flt4 receptor tyrosine kinase ligand and related nucleic acid - used
XX to modulate growth of endothelial cells and for diagnosis of
XX endothelial cell diseases
XX
XX Claim 7; Page 112-113; 183pp; English.
XX
XX This polypeptide comprises the pre-pro sequence of human VEGF-C,
XX a novel ligand that binds specifically to human Flt4 receptor
XX tyrosine kinase (VEGFR-3), stimulating phosphorylation of the
XX receptor. Its sequence was deduced from a cDNA clone (AAT84276)
XX obtd. from a PC-3 prostatic adenocarcinoma cell (ATCC CRL 1435)
XX library. The polypeptide, or its active fragments, can be

PF 02-FEB-1998; 98WO-US01973.
 XX
 PR 05-FEB-1997; 97US-0795430.
 XX
 PA (LUDM-) LUDMIG INST CANCER RES.
 PA (UYHE-) UNIV HELSINKI LICENSING LTD.
 XX
 PI Altalo K, Joukov V;
 XX
 DR WPI; 1998-437470/37.
 XX
 PT New isolated vascular endothelial growth factor polypeptide(s) -
 PT used to develop products for treating, e.g. cancers, inflammation,
 PT oedema, granulocytopenia or for wound healing or tissue
 PT transplantation
 XX
 PS Example 35; Page 143-145; 177pp; English.
 XX
 CC The vascular endothelial growth factor C (VEGF-C) polypeptides have
 CC activities affecting growth and migration of vascular endothelial cells,
 CC promoting growth of lymphatic endothelial cells and lymphatic vessels,
 CC increasing vascular permeability, and affecting myelopoiesis. The
 CC products can be used for stimulating angiogenesis, for inhibiting
 CC angiogenesis, for stimulating lymphangiogenesis, treatment or prevention
 CC of inflammation, oedema, elephantiasis, or Milroy's disease. They can
 CC also be used to modulate myelopoiesis, e.g. treating granulocytopenia.
 CC They can also be used for modulating the growth of endothelial cells.
 CC They can also be used to stimulate lymphocyte production and maturation,
 CC and to promote or inhibit trafficking of leucocytes between tissues and
 CC lymphatic vessels or to affect migration in and out of the thymus.
 CC
 SQ Sequence 419 AA;

Query Match 100.0%; Score 1130; DB 19; Length 419;
 Best Local Similarity 100.0%; Pred. No. 3.1e-82;
 Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLPATLPOCAANKTCPTNYMNMNHICRCLAOEDMFSSDADSDSTDFHDCGNKELD 60
 |||||||
 DB 228 slpatlpqgaanktcptnywmnhicrclaqedfmfssdagdstdfhdcgpnkeld 287
 |||||||
 OY 61 EETCCVCVCRAGLRPASCGPHKELDRNSCCQVCVCKNLFPSCGCGANREPDENTCCVCCKRTC 120
 |||||||
 DB 288 eeltcgcvcraglrpasgphkeldrnsccqvcvcknlfpsgcganrefdentcgcvcckrtc 347
 |||||||
 OY 121 PRNOPLNPGKCACTESPOKCLLGKKFHHOTGSCYRPPCNRQKACPEPGSYSEYVCR 180
 |||||||
 DB 348 prngplnpgkcactespgkcllgkffhgtcscyrppctnrqkacepgfsyseevcr 407
 |||||||
 OY 181 CVPSYWKRPOMS 192
 |||||||
 DB 408 cvpsywkrpoms 419

RESULT 11
 AAM75740
 ID AAM75740 standard; Protein; 419 AA.

AC AAM75740;
 XX
 DT 20-NOV-1998 (first entry)
 XX
 DE Human vascular endothelial growth factor C protein.
 XX
 KM Flt4; vascular endothelial growth factor C; vascular endothelial cell;
 KM lymphatic endothelial cell; myelopoiesis; angiogenesis; inflammation;
 KM lymphangiogenesis; oedema; elephantiasis; Milroy's disease.
 XX
 OS Homo sapiens.
 XX
 PN MO9833917-A1.
 XX

PD 06-AUG-1998.
 XX
 PF 02-FEB-1998; 98WO-US01973.
 XX
 PR 05-FEB-1997; 97US-0795430.
 XX
 PA (LUDM-) LUDMIG INST CANCER RES.
 PA (UYHE-) UNIV HELSINKI LICENSING LTD.
 XX
 PI Altalo K, Joukov V;
 XX
 DR WPI; 1998-437470/37.
 DR N-PSDB; AAV52576.
 XX
 PT New isolated vascular endothelial growth factor polypeptide(s) -
 PT used to develop products for treating, e.g. cancers, inflammation,
 PT oedema, granulocytopenia or for wound healing or tissue
 PT transplantation
 XX
 PS Claim 1; Page 112-115; 177pp; English.
 XX
 CC The vascular endothelial growth factor C (VEGF-C) polypeptides have
 CC activities affecting growth and migration of vascular endothelial cells,
 CC promoting growth of lymphatic endothelial cells and lymphatic vessels,
 CC increasing vascular permeability, and affecting myelopoiesis. The
 CC products can be used for stimulating angiogenesis, for inhibiting
 CC angiogenesis, for stimulating lymphangiogenesis, treatment or prevention
 CC of inflammation, oedema, elephantiasis, or Milroy's disease. They can
 CC also be used to modulate myelopoiesis, e.g. treating granulocytopenia.
 CC They can also be used for modulating the growth of endothelial cells.
 CC They can also be used to stimulate lymphocyte production and maturation,
 CC and to promote or inhibit trafficking of leucocytes between tissues and
 CC lymphatic vessels or to affect migration in and out of the thymus.
 CC
 SQ Sequence 419 AA;

Query Match 100.0%; Score 1130; DB 19; Length 419;
 Best Local Similarity 100.0%; Pred. No. 3.1e-82;
 Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLPATLPOCAANKTCPTNYMNMNHICRCLAOEDMFSSDADSDSTDFHDCGNKELD 60
 |||||||
 DB 228 slpatlpqgaanktcptnywmnhicrclaqedfmfssdagdstdfhdcgpnkeld 287
 |||||||
 OY 61 EETCCVCVCRAGLRPASCGPHKELDRNSCCQVCVCKNLFPSCGCGANREPDENTCCVCCKRTC 120
 |||||||
 DB 288 eeltcgcvcraglrpasgphkeldrnsccqvcvcknlfpsgcganrefdentcgcvcckrtc 347
 |||||||
 OY 121 PRNOPLNPGKCACTESPOKCLLGKKFHHOTGSCYRPPCNRQKACPEPGSYSEYVCR 180
 |||||||
 DB 348 prngplnpgkcactespgkcllgkffhgtcscyrppctnrqkacepgfsyseevcr 407
 |||||||
 OY 181 CVPSYWKRPOMS 192
 |||||||
 DB 408 cvpsywkrpoms 419

RESULT 12
 AAM86203
 ID AAM86203 standard; protein; 419 AA.

AC AAM86203;
 XX
 DT 16-FEB-1999 (first entry)
 XX
 DE Human vascular endothelial growth factor (VEGF)-C sequence.
 XX
 KM VEGF; VRF; vascular endothelial growth factor; VEGF-related protein;
 KM recombinant; truncated; gene therapy; angiogenesis; cardiac ischemia;
 KM coronary; collateral vessel development; cell growth; migration; heart;
 KM lower limb ischemia; stroke; peripheral vascular disease; intestine;
 KM wound healing; skin; vascular permeability.

```

XX OS Homo sapiens.
XX XX MO9849300-A2.
XX PD 05-NOV-1998.
XX PF 20-APR-1998; 98WO-US07801.
XX PR 25-APR-1997; 97US-0842984.
XX PA (COLL-) COLLATERAL THERAPEUTICS.
XX PI Bohlén P;
XX DR WPI; 1999-009426/01.
XX PT New truncated vascular endothelial growth factor-related protein
XX PT subunits - lack part of the N-terminal sequence, used to stimulate
XX PT angiogenesis, e.g. for treating heart disease and ischemia
XX PS Disclosure: Fig 1; 113pp; English.
XX CC This represents the amino acid sequence of human vascular endothelial
XX CC growth factor (VEGF)-C protein. The invention provides truncated VRF
XX CC (VEGF-related protein) subunits that have at least one amino acid
XX CC N-terminal to the first Cys of the core sequence deleted. Host cells
XX CC transformed or transfected with expression vectors containing nucleic
XX CC acids encoding the truncated VRF subunits are used to produce the
XX CC truncated proteins recombinantly. The truncated VRF subunits, optionally
XX CC expressed from gene therapy vectors, have in vivo and in vitro angiogenic
XX CC activity and are used to stimulate angiogenesis, particularly coronary
XX CC collateral vessel development in cases of cardiac ischemia; to stimulate
XX CC endothelial cell growth and migration in vitro; to treat heart disease;
XX CC to treat ischemia (e.g. cardiac, chronic coronary or chronic lower limb
XX CC ischemia; stroke and peripheral vascular disease); to promote healing of
XX CC wounds (of skin or intestines), and to increase vascular permeability.
XX SQ Sequence 419 AA;

Query Match 100.0%; Score 1130; DB 20; Length 419;
Best Local Similarity 100.0%; Pred. No. 3.1e-82;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLPATLPOQAANKTCPTNYMNNHICRCLAQEDMFSSDAGDSDTDFHIDCGPKKELD 60
DB 228 slpatlpqgaanktcptnymnnhicrlaqedfmfssdagdsdtgfhidicgpnkeld 287
QY 61 EETCOCVCGAGLRPASCGPHKELDRNSCQVCYCNKLFPSQCGANREPDENTCQVCCKRTC 120
DB 288 eetcgcvctaglrpasgphkeldrnsqgcvcvknklfsgcganrefdentcgvcckrtc 347
QY 121 PRNOPLNPGKCAKCECTESFOKCLLGKKFHHOTGSCYRRPCTNNRKACBPGFSYSEVCR 180
DB 348 prngplnpgkcaectespqckllgkxfhqtcsyrrpctnrqacbpfsyseevcr 407
QY 181 CVPSTYKRRPOMS 192
DB 408 cvpsywkrrpoms 419

RESULT 13
AAB10648
ID AAB10648 standard; Protein: 419 AA.
XX AC AAB10648;
XX AC 19-JAN-2001 (first entry)
XX DE Human VEGC protein.
XX KW VEGF-X; vascular endothelial growth factor; human; vulnery; cytostatic;

```

```

KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
KW venous sore; diabetic ulcer; burns; skin graft growth; VEGC.
XX OS Homo sapiens.
XX PN MO200037641-A2.
XX PD 29-JUN-2000.
XX PF 21-DEC-1999; 99WO-US30503.
XX PR 22-DEC-1998; 98GB-0028377.
XX PR 18-MAR-1999; 99US-0124967.
XX PR 08-NOV-1999; 99US-0164131.
XX PA (JANC ) JANSSEN PHARM NV.
XX PI Gordon RD, Sprengel JJ, Yon JR, Dijkmans JH, Gosielska A;
XX PI Dhanraj SN, Xu J;
XX DR WPI; 2000-442669/38.
XX PT New vascular endothelial growth factor protein, useful for treating or
XX PT preventing diseases associated with inappropriate angiogenesis activity
XX PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
XX PS Disclosure: Fig 11; 127pp; English.
XX CC This invention describes a novel vascular endothelial growth factor-X
XX CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
XX CC vulnery, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
XX CC antidiabetic activity and acts as an angiogenesis and vascularization
XX CC regulator. An antisense molecule of the invention is useful for treating
XX CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
XX CC retinopathy by inhibiting angiogenic activity or inappropriate
XX CC vascularization including formation and proliferation of new blood
XX CC vessels, growth and development of tissues, tissue regeneration and organ
XX CC and tissue repair in a subject. The products of the invention are useful
XX CC for preparing medicaments for treating wounds such as dermal ulcers,
XX CC pressure sores, venous sores, diabetic ulcers and burns and to promote
XX CC skin graft growth, tissue repair, proliferation of new blood vessels,
XX CC tissue regeneration and organ repair by promoting angiogenic activity or
XX CC vascularization. This sequence represents the human VEGC protein used
XX CC to illustrate the method of the invention.
XX SQ Sequence 419 AA;

Query Match 100.0%; Score 1130; DB 21; Length 419;
Best Local Similarity 100.0%; Pred. No. 3.1e-82;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLPATLPOQAANKTCPTNYMNNHICRCLAQEDMFSSDAGDSDTDFHIDCGPKKELD 60
DB 228 slpatlpqgaanktcptnymnnhicrlaqedfmfssdagdsdtgfhidicgpnkeld 287
QY 61 EETCOCVCGAGLRPASCGPHKELDRNSCQVCYCNKLFPSQCGANREPDENTCQVCCKRTC 120
DB 288 eetcgcvctaglrpasgphkeldrnsqgcvcvknklfsgcganrefdentcgvcckrtc 347
QY 121 PRNOPLNPGKCAKCECTESFOKCLLGKKFHHOTGSCYRRPCTNNRKACBPGFSYSEVCR 180
DB 348 prngplnpgkcaectespqckllgkxfhqtcsyrrpctnrqacbpfsyseevcr 407
QY 181 CVPSTYKRRPOMS 192
DB 408 cvpsywkrrpoms 419

RESULT 14

```


xx	AAAB29048	
ID	AAAB29048 standard; Protein; 419 AA.	
xx		
AC	AAAB29048;	
xx		
DT	31-JAN-2001 (first entry)	
xx		
DE	Human VEGF-C protein sequence.	
xx		
KW	Human; Flt4; fms-like tyrosine kinase 4; lymphoedema;	
KW	vascular endothelial growth factor receptor 3; VEGFR-3;	
KW	Mitroy-Nonne syndrome; lymphoedema praecox; VEGF-C;	
KW	vascular endothelial growth factor C.	
xx		
OS	Homo sapiens.	
xx		
PN	WO200058511-A1.	
xx		
PD	05-OCT-2000.	
xx		
PF	26-MAR-1999; 99WO-US06133.	
xx		
PR	26-MAR-1999; 99WO-US06133.	
xx		
PA	(LUDWG-) LUDWIG INST CANCER RES.	
PA	(UYHE-) UNIV HELSINKI LICENSING LTD OY.	
PA	(UYPL-) UNIV PITTSBURGH.	
xx		
PI	Ferrell RE, Alltalo K, Finegold DN, Karkkainen M;	
xx		
DR	WPI: 2000-679298/66.	
DR	N-PSDB; AAC62406.	
xx		
PT	Screening a human subject for increased risk of developing a lymphatic	
PT	disorder, comprises assaying a nucleic acid to determine a mutation	
PT	altering the sequence of a vascular endothelial growth factor	
PT	receptor-3 -	
xx		
PS	Disclosure: Page 60-61; 76pp; English.	
xx		
CC	The present sequence is the protein sequence for the human vascular	
CC	endothelial growth factor C (VEGF-C). It was used to demonstrate the	
CC	methods of the invention, which involve the screening of individuals to	
CC	determine which vascular endothelial growth factor receptor 3 (VEGFR-3,	
CC	also known as Flt4 or fms-like tyrosine kinase 4) alleles they possess	
CC	and thus their likelihood of developing hereditary lymphoedema.	
CC	Conditions associated with lymphoedema include Mitroy-Nonne syndrome,	
CC	which is early onset lymphoedema and lymphoedema praecox, which is late	
CC	onset.	
xx		
Sequence	419 AA;	
xx		

Query Match	100.0%	Score 1130	DB 21	Length 419
Best Local Similarity	100.0%	Prod. NO. 3.1e-82		
Matches 192	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 1	SLPATLPOCQAANKTCPTNYMWNHHCRCALQEDFMFSDDAGDDSTDFGHDICGNKELD	60		
Db	slpatlpqgqaanktcptnymwnhlcrcalagedfmfsdagddstdfghdicgnkeld	287		
QY 61	EETOCVCVRAGIRPASCGRPHKELDNSOCVCYKNTLFPSCGANPEFENOCVCYCKRRC	120		
Db	eetccvcvcraglrpascgrphkeldnscqcycknllfsgqeanrefdentccvckrrc	347		
QY 121	PRNOLPBGKCAECETESPOKCLTGKRFHNOTSCYRRPCTNRKACBPGFSYSEEVCR	180		
Db	prnplpbgkcaecetespqckllkgkrfhnbftscyrtrpctnrkacbpgrfsyseevcr	407		
QY 181	CVPSYWKRRPQMS	192		
Db	cvpsywkkrrpms	419		

CC	XX	DE	Human prepro-vascular endothelial growth factor C.
CC	XX	DE	Human; receptor tyrosine kinase; RTK; Flt4; fms-like tyrosine kinase 4;
CC	KW	KW	VEGFR-3; vascular endothelial growth factor receptor-3; chromosome 5q35;
CC	KW	KW	cystosatic; tumour imaging; anti-tumour therapy; treatment; diagnosis;
CC	KW	KW	neoplastic disease; lymphoma; carcinoma; breast; squamous cell; melanoma;
CC	KW	KW	sarcoma; malignancy; VEGF-C; vascular endothelial growth factor C.
CC	OS	OS	Homo sapiens.
CC	FH	FH	Location/Qualifiers
CC	FT	FT	1..31
CC	FT	FT	/label= Signal_peptide
CC	FT	FT	32..103
CC	FT	FT	/label= N-terminal_peptide
CC	FT	FT	/note= "cleavage of this peptide from partially processed
CC	FT	FT	VEGF-C produces a fully processed mature form of VEGF-C
CC	FT	FT	of 21-23 kD which has high affinity to VEGFR-2"
CC	FT	FT	104..227
CC	FT	FT	/label= Mature_VEGF-C
CC	FT	FT	228..419
CC	FT	FT	/label= C-terminal_peptide
CC	FT	FT	/note= "has a pattern of spaced cysteine residues
CC	FT	FT	reminiscent of a Balbian ring 3 protein (BR3p) sequence;
CC	FT	FT	cleavage of signal peptide and the C-terminal
CC	FT	FT	peptide produces a partially processed form of VEGF-C of
CC	FT	FT	about 29 kD which has high affinity to Flt4 (VEGFR-3)"
CC	FT	FT	113..213
CC	FT	FT	/note= "binds and stimulates VEGF-C receptors; Cys
CC	FT	FT	at position 156 is essential for VEGFR-2 binding and at
CC	FT	FT	156 is essential for VEGFR-2 and VEGFR-3 binding"
CC	FT	FT	131..211
CC	FT	FT	/note= "important for VEGF-C activity"
CC	FT	FT	Region
CC	PN	PN	MO200021560-A1.
CC	PD	PD	20-Apr-2000.
CC	XX	XX	08-Oct-1999; 99WO-US23525.
CC	XX	XX	09-Oct-1998; 98US-0169079.
CC	PA	PA	(LUDM-) LUDMIG INST CANCER RES.
CC	PA	PA	(UTHE-) UNIV HELSINKI LICENSING LTD OY.
CC	P1	P1	Alltalo K, Kaipainen A, Valtola R, Jussila I;
CC	DR	DR	WPI; 2000-317850/27.
CC	PT	PT	Treating neoplastic diseases such as lymphoma, carcinomas, melanomas
CC	PT	PT	and sarcomas, involves administering a compound capable of inhibiting
CC	PT	PT	binding of ligand/proteins to fms-like tyrosine kinase-4 receptor -
CC	PS	PS	Example 15-17; Page 140-142; 148pp: English.
CC	XX	XX	The patent discloses a method to treat neoplastic-disease characterised
CC	CC	CC	by expression of fms-like tyrosine kinase 4 (Flt4) receptor (also
CC	CC	CC	referred as vascular endothelial growth factor receptor-3, VEGFR-3) in
CC	CC	CC	endothelial cells of blood vessels adjacent to malignant neoplasm. The
CC	CC	CC	method involves administering a compound that inhibits binding of a
CC	CC	CC	ligand to Flt4 thereby inhibiting Flt4 mediated proliferation of vascular
CC	CC	CC	endothelial cells. The compound is useful for treating neoplastic disease
CC	CC	CC	such as breast carcinomas, squamous cell carcinomas, lymphomas, melanomas
CC	CC	CC	and sarcomas. Flt4 receptor tyrosine kinase binding compounds can be used

CC for manufacturing medicament useful for diagnostic screening, imaging and
CC treatment of malignancies characterised by Flt4-expressing blood cells.
CC The Flt4 gene maps to chromosomal region 5q35 and is expressed as 5.8 kb
CC and 4.5 kb mRNAs which differ in their 3' sequences and are
CC differentially expressed in HEL and DAMI cell lines. Flt4
CC belongs to a subfamily of class III receptor tyrosine kinases (RTKs).
CC It is used as a target for tumour imaging and anti-tumour therapy.
CC The present sequence is a human prepro-vascular endothelial growth
CC factor C (VEGF-C), a specific example of Flt4 binding compound.
XX
SQ Sequence 419 AA:

Query Match 100.0%; Score 1130; DB 21; Length 419;
Best Local Similarity 100.0%; Pred. No. 3.1e-82;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SLPATLPQCQAAANKTCPTNYMMNNHICRCLAOEDFMFSSDAGDDSTDGFHDICGPNKELD 60
DB 228 sipatlpqqaanktcptnymmnnhircrlagedfmfssdagddstdgfhdlcgpnkeld 287
OY 61 EETCOCVCAGLRPASCGPHKELDRNSCOCVCCKNLFPSCGANREPDENTCOCVCKRTG 120
DB 288 eetcocvcaglrpascgphkeldrnsccvccknlfpscganrefdentcvcckrtg 347
OY 121 PRNOPLPNGKACECTESPQKCLLKGGKPFHOTCSCYRRPCTNRKACEPGFSYSEVCR 180
DB 348 prngplngpkacectespqckllkgykfhqtcscyrpctnrqkacepgfsysevcr 407
OY 181 CVPSYWKRPQMS 192
DB 408 cvpsywkrrpqms 419

Search completed: October 17, 2001, 14:46:01
Job time: 127 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:47:10 ; Search time 62.93 seconds
(without alignments)
62.821 Million cell updates/sec

Title: US-09-427-657-2_COPY_228_419

Sequence: 1 SLPATLPQCAANKTCPTNY.....SYSEVCRVPSYWRPQMS 192

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents-AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1130	100.0	350	4	US-08-510-133A-33 Sequence 33, Appl
2	1130	100.0	350	4	US-08-585-895-33 Sequence 33, Appl
3	1130	100.0	419	4	US-08-795-430-8 Sequence 8, Appl
4	1130	100.0	419	4	US-08-510-133A-35 Sequence 35, Appl
5	1126	99.6	350	2	US-08-999-811-4 Sequence 4, Appl
6	1126	99.6	350	2	US-08-824-996-2 Sequence 2, Appl
7	1126	99.6	350	3	US-09-042-105-4 Sequence 4, Appl
8	1126	99.6	419	3	US-08-999-811-2 Sequence 2, Appl
9	1126	99.6	419	3	US-09-042-105-2 Sequence 2, Appl
10	1126	99.6	419	3	US-09-042-105-18 Sequence 18, Appl
11	1126	99.6	419	5	PCT-US96-09001-2 Sequence 2, Appl
12	998	88.3	415	4	US-08-795-430-11 Sequence 11, Appl
13	839	74.2	418	4	US-08-795-430-13 Sequence 13, Appl
14	234	20.7	325	4	US-08-915-795-3 Sequence 5, Appl
15	234	20.7	325	4	US-08-915-795-5 Sequence 5, Appl
16	225	19.9	358	4	US-08-915-795-8 Sequence 8, Appl
17	189	16.7	321	4	US-08-915-795-9 Sequence 9, Appl
18	151	13.4	1810	5	PCT-US95-11684-4 Sequence 4, Appl
19	143.5	12.7	2813	3	US-08-896-449A-2 Sequence 2, Appl
20	143.5	12.7	2813	3	US-09-132-652-2 Sequence 2, Appl
21	139.5	12.3	2471	1	US-08-185-432-16 Sequence 16, Appl
22	139.5	12.3	2471	1	US-08-083-590A-19 Sequence 19, Appl
23	139.5	12.3	2471	3	US-08-532-384-19 Sequence 19, Appl
24	137.5	12.2	2523	1	US-08-185-432-18 Sequence 18, Appl
25	134	11.9	1219	4	US-08-882-046-5 Sequence 5, Appl
26	133	11.8	642	3	US-08-872-855-10 Sequence 10, Appl
27	130	11.5	1010	4	US-08-882-046-7 Sequence 7, Appl

28	130	11.5	1218	2	US-08-400-159-6 Sequence 6, Appl
29	130	11.5	1218	3	US-08-611-729A-6 Sequence 6, Appl
30	130	11.5	1218	4	US-08-882-046-2 Sequence 2, Appl
31	130	11.5	2199	5	PCT-US95-11684-2 Sequence 2, Appl
32	128	11.3	2703	1	US-08-185-432-19 Sequence 19, Appl
33	125.5	11.1	374	2	US-08-820-170A-25 Sequence 25, Appl
34	125.5	11.1	374	3	US-09-055-699-25 Sequence 25, Appl
35	125.5	11.1	374	4	US-09-273-565-25 Sequence 25, Appl
36	125.5	11.1	374	5	PCT-US95-06385-2 Sequence 2, Appl
37	125	11.1	2556	1	US-08-083-590A-20 Sequence 20, Appl
38	125	11.1	2556	3	US-08-532-384-20 Sequence 20, Appl
39	124	11.0	314	2	US-08-460-309-19 Sequence 19, Appl
40	124	11.0	314	2	US-08-125-077-19 Sequence 19, Appl
41	123	10.9	2050	2	US-08-347-594A-2 Sequence 2, Appl
42	122	10.8	1193	2	US-08-400-159-10 Sequence 10, Appl
43	122	10.8	1193	3	US-08-611-729A-10 Sequence 10, Appl
44	122	10.8	2556	1	US-08-185-432-17 Sequence 17, Appl
45	120	10.6	1248	4	US-08-882-046-6 Sequence 6, Appl

ALIGNMENTS

```
RESULT 1
US-08-510-133A-33
Sequence 33, Application US/08510133A
Patent No. 6221839
GENERAL INFORMATION:
APPLICANT: Joukoy, Vladimir
TITLE OF INVENTION: Receptor Ligand
NUMBER OF SEQUENCES: 35
CORRESPONDING ADDRESSES:
ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/510,133A
FILING DATE: 01-Aug-1995
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28113/32863
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856.
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-08-510-133A-33
Query Match 100.0%; Score 1130; DB 4; Length 350;
Best Local Similarity 100.0%; Pred. No. 3e-95;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SLPATLPQCAANKTCPTNYMNNHICRLAEDFMFSSDAGDSDTGFHDICGNKEID 60
DB 159 SLPATLPQCAANKTCPTNYMNNHICRLAEDFMFSSDAGDSDTGFHDICGNKEID 218
```

QY 61 EETCOCVCAGLRPASCGPHKELDRNSCOCVCKNKLFPSCGANGREPDENTCOCVCKRTC 120
| | | | |
DB 219 EETCOCVCAGLRPASCGPHKELDRNSCOCVCKNKLFPSCGANGREPDENTCOCVCKRTC 278
| | | | |
QY 121 PRNOPLPNGKCAECTESPQKCLLKGRKPHQOTCSQYRRPCTNRKACBPGFSYSEVCR 180
| | | | |
DB 279 PRNOPLPNGKCAECTESPQKCLLKGRKPHQOTCSQYRRPCTNRKACBPGFSYSEVCR 338
| | | | |
QY 181 CVPSTYWKRPQMS 192
| | | | |
DB 339 CVPSTYWKRPQMS 350
| | | | |
RESULT 2
US-08-585-895-33
; Sequence 33, Application US/08585895
; Patent No. 6245530
; GENERAL INFORMATION:
; APPLICANT: Alltalo, Karl
; APPLICANT: Joukov, Vladimir
; TITLE OF INVENTION: Receptor Ligand
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/585,895
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 28113/33072
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 350 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-585-895-33

Query Match 100.0%; Score 1130; DB 4; Length 350;
Best Local Similarity 100.0%; Pred. No. 3e-95;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLPATLPQCAANKTCPTNYMNNHICRCLAOEDFMFSSDAGDDSTDPGHDICGPNKELD 60
| | | | |
DB 159 SLPATLPQCAANKTCPTNYMNNHICRCLAOEDFMFSSDAGDDSTDPGHDICGPNKELD 218
| | | | |
QY 61 EETCOCVCAGLRPASCGPHKELDRNSCOCVCKNKLFPSCGANGREPDENTCOCVCKRTC 120
| | | | |
DB 219 EETCOCVCAGLRPASCGPHKELDRNSCOCVCKNKLFPSCGANGREPDENTCOCVCKRTC 278
| | | | |
QY 121 PRNOPLPNGKCAECTESPQKCLLKGRKPHQOTCSQYRRPCTNRKACBPGFSYSEVCR 180
| | | | |
DB 279 PRNOPLPNGKCAECTESPQKCLLKGRKPHQOTCSQYRRPCTNRKACBPGFSYSEVCR 338
| | | | |
QY 181 CVPSTYWKRPQMS 192
| | | | |
DB 339 CVPSTYWKRPQMS 350
| | | | |

DB 339 CVPSTYWKRPQMS 350
| | | | |
RESULT 3
US-08-795-430-8
; Sequence 8, Application US/08795430
; Patent No. 6130071
; GENERAL INFORMATION:
; APPLICANT: Alltalo, Karl
; APPLICANT: Joukov, Vladimir
; TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
; TITLE OF INVENTION: Protein and Gene, Mutants Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,430
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/ET96/00427
; FILING DATE: 01-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/671,573
; FILING DATE: 28-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/601,132
; FILING DATE: 14-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/585,895
; FILING DATE: 12-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/510,133
; FILING DATE: 01-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,011
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 28967/33691
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-795-430-8

Query Match 100.0%; Score 1130; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 3.6e-95;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLPATLPQCAANKTCPTNYMNNHICRCLAOEDFMFSSDAGDDSTDPGHDICGPNKELD 60
| | | | |
DB 228 SLPATLPQCAANKTCPTNYMNNHICRCLAOEDFMFSSDAGDDSTDPGHDICGPNKELD 287
| | | | |

0y	61	ETCQCVCRAGIRPASC	CPHELDNRNSOCQCKKLP	PSOCGARREDEMTQCVKRTC	120
Db	288	ETCQCVCRAGIRPASC	CPHELDNRNSOCQCKKLP	PSOCGARREDEMTQCVKRTC	347
0y	121	PRNDPLNGKCAECETES	POKCLLKGKKFHQTC	SCYRRPCTNRNOKACEPFSYSEVCR	180
Db	348	PRNDPLNGKCAECETES	POKCLLKGKKFHQTC	SCYRRPCTNRNOKACEPFSYSEVCR	407
0y	181	CVPSTYWRPQMS	192		
Db	408	CVPSTYWRPQMS	419		

```

1      RESULT      4
2      US-08-510-133A-35
3      Sequence 35, Application US/08510133A
4      Patent No. 6221839
5      GENERAL INFORMATION:
6      APPLICANT: Alltalo, Karl
7      Joukov, Vladimir
8      TITLE OF INVENTION: Receptor Ligand
9      NUMBER OF SEQUENCES: 35
10     CORRESPONDENCE ADDRESS:
11     ADDRESSSEE: Marshall, O'Toole, Gerstein, Murray & Borun
12     STREET: 6300 Sears Tower, 233 South Wacker Drive
13     CITY: Chicago
14     STATE: Illinois
15     COUNTRY: United States of America
16     ZIP: 60606-6402
17
18     COMPUTER READABLE FORM:
19     MEDIUM TYPE: Floppy disk
20     COMPUTER: IBM PC compatible
21     OPERATING SYSTEM: PC-DOS/MS-DOS
22     SOFTWARE: PatentIn Release #1.0, Version #1.25
23
24     CURRENT APPLICATION DATA:
25     APPLICATION NUMBER: US/08/510,133A
26     FILING DATE: 01-Aug-1995
27     CLASSIFICATION: <Unknown>
28     ATTORNEY/AGENT INFORMATION:
29     NAME: Gass, David A.
30     REGISTRATION NUMBER: 38,153
31     REFERENCE/DOCKET NUMBER: 28113/32863
32     TELECOMMUNICATION INFORMATION:
33     TELEPHONE: 312/474-6300
34     TELEFAX: 312/474-0448
35     TELEX: 25-3856
36
37     INFORMATION FOR SEQ ID NO: 35:
38     SEQUENCE CHARACTERISTICS:
39     LENGTH: 419 amino acids
40     TYPE: amino acid
41     TOPOLOGY: linear
42     MOLECULE TYPE: protein
43     SEQUENCE DESCRIPTION: SEQ ID NO: 35:
44
45     US-08-510-133A-35
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
94
```

Db 408 CVPSTWKRPQMS 419

```

RESULT      5
US-08-999-811-4
: Sequence 4, Application US/08999811
: Patent No. 5932540
:
: GENERAL INFORMATION:
: APPLICANT: HU, JING-SHAN
: APPLICANT: ROSEN, CRAIG A.
: APPLICANT: CAO, LIANG
: TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
: NUMBER OF SEQUENCE ADDRESSES: 15
:
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
: STREET: 1100 NEW YORK AVENUE
: CITY: WASHINGTON
: STATE: DC
: COUNTRY: USA
: ZIP: 20005
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/999,811
: FILING DATE: HERMITH
: CLASSIFICATION:
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 08/207,550
: FILING DATE: 8-MAR-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/465,968
: FILING DATE: 06-JUN-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: MARKOWICZ, KAREN R.
: REGISTRATION NUMBER: 36,351
: REFERENCE/DOCKET NUMBER: 1488.1000004
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)371-2600
: TELEFAX: (202)371-2540
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 350 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
US-08-999-811-4

```

RESULT 6
US-08-824-996-2
Sequence 2, Application US/08824996B
Patent No. 5935820
GENERAL INFORMATION:
APPLICANT: Hu, Jing-Shan
APPLICANT: Rosen, Craig A.
APPLICANT: Cao, Liang
TITLE OF INVENTION: Polynucleotides Encoding Vascular Endothelial Growth
FILE OF INVENTION: Factor 2
FILE REFERENCE: P112D1
CURRENT APPLICATION NUMBER: US/08/824,996B
CURRENT FILING DATE: 1997-03-27
EARLIER APPLICATION NUMBER: 08/207,550
EARLIER FILING DATE: 1994-03-08
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 350
TYPE: PRT
ORGANISM: Homo sapiens
US-08-824-996-2

Query Match 99.6%; Score 1126; DB 2; Length 350;
Best Local Similarity 99.5%; Pred. No. 6,9e-95;
Matches 191; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLPATLPQOAAKCTPTVMNNHICRCLAOEDFMFSSDAGDDSTDFHDIICGPKKELD 60
DB 159 SLPATLPQOAAKCTPTVMNNHICRCLAOEDFMFSSDAGDDSTDFHDIICGPKKELD 218
QY 61 EETCCVCAGAGLRPASCGRHKEIDRNSCOCVCKNKLFPSCCGANREFDENTCOCVCKRTG 120
DB 219 EETCCVCAGAGLRPASCGRHKEIDRNSCOCVCKNKLFPSCCGANREFDENTCOCVCKRTG 278
QY 121 PRNOPLPKCKACECTESPQCKLLKGRFHQTCSCYRRPCTNRKACBPGEFSYSEVCR 180
DB 279 PRNOPLPKCKACECTESPQCKLLKGRFHQTCSCYRRPCTNRKACBPGEFSYSEVCR 338
QY 181 CVPSTWQRPOMS 192
DB 339 CVPSTWQRPOMS 350

RESULT 7
US-09-042-105-4
Sequence 4, Application US/09042105
Patent No. 6040157
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESSES:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,105
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550

FILING DATE: 8-MAR-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TO BE ASSIGNED
FILING DATE: 24-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: ERIC K. STERNE
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-042-105-4

Query Match 99.6%; Score 1126; DB 3; Length 350;
Best Local Similarity 99.5%; Pred. No. 6,9e-95;
Matches 191; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLPATLPQOAAKCTPTVMNNHICRCLAOEDFMFSSDAGDDSTDFHDIICGPKKELD 60
DB 159 SLPATLPQOAAKCTPTVMNNHICRCLAOEDFMFSSDAGDDSTDFHDIICGPKKELD 218
QY 61 EETCCVCAGAGLRPASCGRHKEIDRNSCOCVCKNKLFPSCCGANREFDENTCOCVCKRTG 120
DB 219 EETCCVCAGAGLRPASCGRHKEIDRNSCOCVCKNKLFPSCCGANREFDENTCOCVCKRTG 278
QY 121 PRNOPLPKCKACECTESPQCKLLKGRFHQTCSCYRRPCTNRKACBPGEFSYSEVCR 180
DB 279 PRNOPLPKCKACECTESPQCKLLKGRFHQTCSCYRRPCTNRKACBPGEFSYSEVCR 338
QY 181 CVPSTWQRPOMS 192
DB 339 CVPSTWQRPOMS 350

RESULT 8
US-08-999-811-2
Sequence 2, Application US/08999811
Patent No. 5932540
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESSES:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/999,811
FILING DATE: HERewith
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550
FILING DATE: 8-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: MARKOWICZ, KAREN R.
REGISTRATION NUMBER: 36,351
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-999-811-2

Query Match 99.6%; Score 1126; DB 2; Length 419;
Best Local Similarity 99.5%; Pred. No. 8.2e-95;
Matches 191; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLPLTPQCOAANKTCPTNYMNNHICRLAODEPMFSSDAGDSDTDGFHDICGPNKELD 60
DB 228 SLPLTPQCOAANKTCPTNYMNNHICRLAODEPMFSSDAGDSDTDGFHDICGPNKELD 287
QY 61 EETCCQCVRAGLRAPASCGPHKELDRNSQCVCYCKNKLFPSCGANEFPDENTCCQCKRRTG 120
DB 288 EETCCQCVRAGLRAPASCGPHKELDRNSQCVCYCKNKLFPSCGANEFPDENTCCQCKRRTG 347
QY 121 PRNOPLNKGKACCECTESPQCKLKGKFFHQTCSYRRPCTNRKACPEPGFSYSEVCR 180
DB 348 PRNOPLNKGKACCECTESPQCKLKGKFFHQTCSYRRPCTNRKACPEPGFSYSEVCR 407
QY 181 CVPSTYWRPOMS 192
DB 408 CVPSTYWRPOMS 419

RESULT 9
US-09-042-105-2
Sequence 2, Application US/09042105
Patent No. 6040157
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: US/09/042,105
APPLICATION NUMBER: US/09/042,105
CLASSIFICATION: HEREWITH
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550
FILING DATE: 8-MAR-1994
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TO BE ASSIGNED
FILING DATE: 24-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: ERIC K. STEFFE
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-042-105-2

Query Match 99.6%; Score 1126; DB 3; Length 419;
Best Local Similarity 99.5%; Pred. No. 8.2e-95;
Matches 191; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLPLTPQCOAANKTCPTNYMNNHICRLAODEPMFSSDAGDSDTDGFHDICGPNKELD 60
DB 228 SLPLTPQCOAANKTCPTNYMNNHICRLAODEPMFSSDAGDSDTDGFHDICGPNKELD 287
QY 61 EETCCQCVRAGLRAPASCGPHKELDRNSQCVCYCKNKLFPSCGANEFPDENTCCQCKRRTG 120
DB 288 EETCCQCVRAGLRAPASCGPHKELDRNSQCVCYCKNKLFPSCGANEFPDENTCCQCKRRTG 347
QY 121 PRNOPLNKGKACCECTESPQCKLKGKFFHQTCSYRRPCTNRKACPEPGFSYSEVCR 180
DB 348 PRNOPLNKGKACCECTESPQCKLKGKFFHQTCSYRRPCTNRKACPEPGFSYSEVCR 407
QY 181 CVPSTYWRPOMS 192
DB 408 CVPSTYWRPOMS 419

RESULT 10
US-09-042-105-18
Sequence 18, Application US/09042105
Patent No. 6040157
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: US/09/042,105
APPLICATION NUMBER: US/09/042,105
CLASSIFICATION: HEREWITH
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550

FILING DATE: 8-MAR-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TO BE ASSIGNED
FILING DATE: 24-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: ERIC K. STEFFE
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-042-105-18

Query Match 99.6% Score 1126; DB 3; Length 419;
Best Local Similarity 99.5% Pred. No. 8.2e-95;

Matches 191; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLPATLPQCAANKTCPTNYMNNHICRCLAQEDFMFSSDAGDSTDGFHDICGPNKELD 60
DB 228 SLPATLPQCAANKTCPTNYMNNHICRCLAQEDFMFSSDAGDSTDGFHDICGPNKELD 287
QY 61 EETCCVCRAGLRPASCGRPHKELDRNSCCVCYCKNKLFPSCGAGNREFEENTCCVCYCKRTC 120
DB 288 EETCCVCRAGLRPASCGRPHKELDRNSCCVCYCKNKLFPSCGAGNREFEENTCCVCYCKRTC 347
QY 121 PRNOPPLNGKACCECTESPOKCLLKGGKHHOTCSYRPPCTNRKACPEPGFSYSEVCR 180
DB 348 PRNOPPLNGKACCECTESPOKCLLKGGKHHOTCSYRPPCTNRKACPEPGFSYSEVCR 407
QY 181 CVPSTYWRPQMS 192
DB 408 CVPSTYWRPQMS 419

RESULT 11
PCT-US96-09001-2
Sequence 2, Application PC/TUS9609001
GENERAL INFORMATION:
APPLICANT: HU, ET AL.
TITLE OF INVENTION: Human Vascular Endothelial Growth Factor 2
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELIA, BYRNE, BAIN, GILFILLAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/09001
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/465,968
FILING DATE: 6 JUN 95
APPLICATION NUMBER: 08/207,550
FILING DATE: 8 MAR 1994
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-288
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
PCT-US96-09001-2

Query Match 99.6% Score 1126; DB 5; Length 419;
Best Local Similarity 99.5% Pred. No. 8.2e-95;

Matches 191; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLPATLPQCAANKTCPTNYMNNHICRCLAQEDFMFSSDAGDSTDGFHDICGPNKELD 60
DB 228 SLPATLPQCAANKTCPTNYMNNHICRCLAQEDFMFSSDAGDSTDGFHDICGPNKELD 287
QY 61 EETCCVCRAGLRPASCGRPHKELDRNSCCVCYCKNKLFPSCGAGNREFEENTCCVCYCKRTC 120
DB 288 EETCCVCRAGLRPASCGRPHKELDRNSCCVCYCKNKLFPSCGAGNREFEENTCCVCYCKRTC 347
QY 121 PRNOPPLNGKACCECTESPOKCLLKGGKHHOTCSYRPPCTNRKACPEPGFSYSEVCR 180
DB 348 PRNOPPLNGKACCECTESPOKCLLKGGKHHOTCSYRPPCTNRKACPEPGFSYSEVCR 407
QY 181 CVPSTYWRPQMS 192
DB 408 CVPSTYWRPQMS 419

RESULT 12
US-08-795-430-11
Sequence 11, Application US/08795430
Patent No. 6130071
GENERAL INFORMATION:
APPLICANT: Alitalo, Kari
APPLICANT: Joukov, Vladimir
TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
TITLE OF INVENTION: Protein and Gene, Mutants Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,430
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FT196/00427
FILING DATE: 01-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/671,573

FLING DATE: 28-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/601,132
FLING DATE: 14-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/585,895
FLING DATE: 12-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/510,133
FLING DATE: 01-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,011
FLING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28967/33691
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 415 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-795-430-11

Query Match 88.3%; Score 998; DB 4; Length 415;
Best Local Similarity 84.4%; Pred. No. 3.2e-83;
Matches 162; Conservative 18; Mismatches 12; Indels 0; Gaps 0;

QY 1 SLPTLPQQAANKTCPTNYMNNHICRCLAOEDFMFSSDADSTIDGFHICGPKNELD 60
DB 224 SLPTLPQQAANKTCPTNYMNNHICRCLAOEDFMFSSDADSTIDGFHICGPKNELD 283
QY 61 EETCCGVCAGLRPASCGRPHKELDRNSCCQVCYCKNKLFPSCGANKREPDENTCQVCCKRT 120
DB 284 EETCCGVCAGLRPASCGRPHKELDRNSCCQVCYCKNKLFPSCGANKREPDENTCQVCCKRT 343
QY 121 PRNOPLPNGKACCECTESPOKCLGKKFHHQTCSCYRRPCTNRQACGEPFSSEYVCR 180
DB 344 PRNOPLPNGKACCECTESPOKCLGKKFHHQTCSCYRRPCTNRQACGEPFSSEYVCR 403
QY 181 CVPSTYWKRPQMS 192
DB 404 CVPSTYWKRPQMS 415

RESULT 13
US-08-795-430-13
Sequence 13, Application US/08795430
Patent No. 6130071
GENERAL INFORMATION:
APPLICANT: Alitalo, Kari
APPLICANT: Joukov, Vladimir
TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,430
FLING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FI96/00427
FLING DATE: 01-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/671,573
FLING DATE: 28-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/510,132
FLING DATE: 14-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/585,895
FLING DATE: 12-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,011
FLING DATE: 01-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,011
FLING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28967/33691
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 418 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-795-430-13

Query Match 74.2%; Score 839; DB 4; Length 418;
Best Local Similarity 73.1%; Pred. No. 8.1e-69;
Matches 141; Conservative 17; Mismatches 33; Indels 2; Gaps 2;

QY 1 SLPTLPQQAANKTCPTNYMNNHICRCLAOEDFMFSSDADSTIDGFHICGPKNEL 59
DB 227 SLPTLPQQAANKTCPTNYMNNHICRCLAOEDFMFSSDADSTIDGFHICGPKNEL 285
QY 60 DEETCCGVCAGLRPASCGRPHKELDRNSCCQVCYCKNKLFPSCGANKREPDENTCQVCCKRT 119
DB 286 DEETCCGVCAGLRPASCGRPHKELDRNSCCQVCYCKNKLFPSCGANKREPDENTCQVCCKRT 345
QY 120 CRRNOPLPNGKACCECTESPOKCLGKKFHHQTCSCYRRPCTNRQACGEPFSSEYVCR 179
DB 346 CRRNOPLPNGKACCECTESPOKCLGKKFHHQTCSCYRRPCTNRQACGEPFSSEYVCR 405
QY 180 RCVPTSTYWKRPQMS 192
DB 406 RCVPTSTYWKRPQMS 418

RESULT 14
US-08-915-795-3
Sequence 3, Application US/08915795
Patent No. 6235713
GENERAL INFORMATION:
APPLICANT: Marc G. ACHEN
APPLICANT: Andrew F. WILKS
APPLICANT: Steven A. STACKER
APPLICANT: Kari ALITALO
TITLE OF INVENTION: GROWTH FACTOR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.

STREET: 1200 G Street, NW, Suite 700
CITY: Washington
STATE: DC
COUNTRY: United States of America
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,795
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/42983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
TELEX: N/A
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 325 amino acids
TYPE: amino acid
STRANDEDNESS: single
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
TISSUE TYPE: Human Breast
US-08-915-795-3

Query Match 20.7%; Score 234; DB 4; Length 325;
Best Local Similarity 28.8%; Pred. No. 3.9e-14;
Matches 51; Conservative 24; Mismatches 46; Indels 56; Gaps 6;

QY 8 OCOANKTCPTNYMNNHICRCLAQEDFMFSSDAGDDSTDFGHDICGPNKLEDEETCCGV 67
DB 165 RSHSKKLCIPIDMLMDSNCKKCVLOE-----ENPLAGED-----HSHLOE----- 225
QY 68 CRAGLRPASCGPHKELDRNSCCVCCKNKLFPSCGANREEDENTCCVCCKRTCPRNOPLN 127
DB 226 -----PALCGPH-----MMFDDRECVCKTCCPCPDLIQH 255
QY 128 PGKAC-ECTESPQCLKGGKFFHQTSC-----YRRPCTNRKACPEPGFSYSEE 177
DB 256 PNCSCFECKESLETCCOKHKLPHDTCSCEDRCFPHRPCASGKTACAKHCRFPKE 312

RESULT 15
US-08-915-795-5
Sequence 5, Application US/08915795
Patent No. 6235713
GENERAL INFORMATION:
APPLICANT: Marc G. ACHEN
APPLICANT: Andrew F. WILKS
APPLICANT: Steven A. STACKER
APPLICANT: Karl ALITALO
TITLE OF INVENTION: GROWTH FACTOR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
STREET: 1200 G Street, NW, Suite 700
CITY: Washington
STATE: DC
COUNTRY: United States of America
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,795
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/42983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
TELEX: N/A
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 354 amino acids
TYPE: amino acid
STRANDEDNESS: single
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
TISSUE TYPE: Human Lung
US-08-915-795-5

Query Match 20.7%; Score 234; DB 4; Length 354;
Best Local Similarity 28.8%; Pred. No. 4.3e-14;
Matches 51; Conservative 24; Mismatches 46; Indels 56; Gaps 6;

QY 8 OCOANKTCPTNYMNNHICRCLAQEDFMFSSDAGDDSTDFGHDICGPNKLEDEETCCGV 67
DB 214 RSHSKKLCIPIDMLMDSNCKKCVLOE-----ENPLAGED-----HSHLOE----- 254
QY 68 CRAGLRPASCGPHKELDRNSCCVCCKNKLFPSCGANREEDENTCCVCCKRTCPRNOPLN 127
DB 255 -----PALCGPH-----MMFDDRECVCKTCCPCPDLIQH 284
QY 128 PGKAC-ECTESPQCLKGGKFFHQTSC-----YRRPCTNRKACPEPGFSYSEE 177
DB 285 PNCSCFECKESLETCCOKHKLPHDTCSCEDRCFPHRPCASGKTACAKHCRFPKE 341

Search completed: October 17, 2001, 14:47:10
Job time: 196 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:48:38 ; Search time 78.16 seconds
(without alignments)
187.123 Million cell updates/sec

Title: US-09-427-657-2_COPY_228_419

Perfect score: 1130

Sequence: 1 SLPATLPQCOAANKTCPTNY.....SYSEEVCRGVSYWKRPPQS 192

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database : PIR_68:*

1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1130	100.0	419	2 S69207	vascular endotheli
2	238	21.1	1700	2 S08167	Balbiani ring 3 pr
3	180.5	16.0	160	2 U00542	18K secretory pro
4	151	13.4	1810	1 A32230	tenascin precursor
5	146.5	13.0	473	2 A56175	adhesive plaque pr
6	139.5	12.3	1203	2 A49175	Notch B protein -
7	139.5	12.3	1746	1 S19694	tenascin precursor
8	139.5	12.3	2195	2 T34264	hypothetical prote
9	139.5	12.3	2471	2 A49128	cell-fate determin
10	138	12.2	647	2 A43902	tenascin - eastern
11	137.5	12.2	2437	2 S42612	transmembrane prot
12	137.5	12.2	2524	2 A35844	Notch protein - Af
13	136	12.0	782	2 A51625	tenascin-like prot
14	136	12.0	2907	2 A57278	fibritillin-2 precu
15	135.5	12.0	1187	2 T18355	hypothetical prote
16	135.5	12.0	4006	2 T09070	probable tenascin
17	135	11.9	2918	2 A54105	fibritillin-2 precu
18	134.5	11.9	1620	2 T27283	hypothetical prote
19	134	11.9	1220	2 A56136	Notch protein pre
20	134	11.9	2352	2 T30201	Notch homolog prot
21	133.5	11.7	565	2 T16408	hypothetical prote
22	132	11.7	2019	1 U01322	tenascin precursor
23	130	11.5	2301	2 A52160	tenascin-C - human
24	128.5	11.4	1372	2 T25933	hypothetical prote
25	128.5	11.4	2871	2 A55624	fibritillin-1 precu
26	128.5	11.4	3002	2 A47221	fibritillin-1 precu
27	128	11.3	2703	1 A24420	notch protein - fr
28	127.5	11.3	2321	2 S78549	notch3 protein - h
29	127	11.2	626	2 T27319	hypothetical prote

30	127	11.2	2871	2 A55567	fibritillin-1 - bovi
31	127	11.2	4135	2 T42629	tenascin-X - bovin
32	126.5	11.2	497	2 T27827	hypothetical prote
33	126	11.2	3635	2 T10053	laminin alpha 5 ch
34	125.5	11.1	346	2 T46914	hypothetical prote
35	125.5	11.1	1111	2 T26972	hypothetical prote
36	125.5	11.1	2531	2 T31070	notch homolog - se
37	124	11.0	570	2 A48836	fibropellin C prec
38	124	11.0	3712	2 S18253	laminin alpha-1 ch
39	123.5	10.9	220	2 S29105	antistatin - Hydra
40	123.5	10.9	1531	2 T42218	slit-1 protein hom
41	123.5	10.9	2555	2 A40043	notch protein homo
42	123.5	10.9	3020	2 A43932	muscle 2 precursor
43	123	10.9	1106	2 T13938	gene shuttle craft
44	123	10.9	1106	2 T44598	hypothetical prote
45	123	10.9	2318	2 S45306	notch 3 protein -

ALIGNMENTS

RESULT 1
S69207
vascular endothelial growth factor C precursor - human
N:Alternate names: FLT4 ligand DHM
C:Species: Homo sapiens (man)
C:Date: 27-Apr-1996 #sequence_revision 01-Nov-1996 #text_change 08-Oct-1999
C:Accession: S69207; S61795; S71443; S69208; G02659
R:Joukov, V.; Pajusola, K.; Kalpalinen, A.; Chilov, D.; Lahtinen, I.; Kuk, E.; Saksel
EMBO J. 15, 1751, 1996
A>Title: Corrigendum: A novel vascular endothelial growth factor, VEGF-C, is a ligand
A:Reference number: S69207; MUID:96203094
A:Accession: S69207
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-419 <JOU>
A:Cross-references: EMBL:X94216; NID:q1177488; PID:CAA63907.1; PID:e221096; PID:g118
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1995
A>Note: only a part of the translation is shown
A>Note: this is a revision to the sequence from reference S61795
R:Joukov, V.; Pajusola, K.; Kalpalinen, A.; Chilov, D.; Lahtinen, I.; Kuk, E.; Saksel
EMBO J. 15, 290-298, 1996
A>Title: A novel vascular endothelial growth factor, VEGF-C, is a ligand for the FLT4
A:Reference number: S61795; MUID:96178224
A:Accession: S61795
A>Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 70-419 <JOU>
A>Note: this sequence has been revised in reference S69207
A:Accession: S71443
A:Molecule type: protein
A:Residues: 1-419 <JOU>
R:Lee, J.; Gray, A.; Yuan, J.; Luoh, S.M.; Avraham, H.; Wood, W.I.
submitted to the EMBL Data Library, December 1995
A:Description: Vascular endothelial growth factor related protein (VRP): A ligand and
A:Reference number: S69208
A:Accession: S69208
A:Molecule type: mRNA
A:Residues: 1-419 <LEE>
A:Cross-references: EMBL:U43142; NID:g1150988; PID:AAA85214.1; PID:g1150989
R:Morris, J.C.
submitted to the EMBL Data Library, May 1996
A:Reference number: H01557
A:Accession: G02659
A>Status: preliminary; translated from GB/EMBL/DDBT
A:Molecule type: mRNA
A:Residues: 1-419 <MOR>
A:Cross-references: EMBL:U58111; NID:g1373426; PID:AA802909.1; PID:g1373427
C:Genetics:
A:Gene: GDB:VBSGC: VRP
A:Cross-references: GDB:3890883; OMIM:601528
F:1-12/Domain: signal sequence #status predicted <SIG>
F:13-102/Domain: propeptide #status predicted <PRO>

A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual trs
A.Molecule type: mRNA
A.Residues: 1-204, 'G', 206-221, 'A', 223-380, 'D', 382-386, 'H', 388-444, 'HN', 447-450, 'V', 452-5
A.Cross-references: GB:M23121
R:Pearson, C.A.; Pearson, D.; Shihabara, S.; Hofsteenge, J.; Chiquet-Ehrismann, R.
EMBO J. 7, 2977-2982, 1988
A.Title: Tenascin: cDNA cloning and induction by TGF-beta.
A.Reference number: S01292; MUID:89030589
A.Accession: S01292
A>Status: not compared with conceptual translation
A.Molecule type: mRNA
A.Residues: 27-181, 'R', 183-204, 'G', 206-221, 'A', 223-380, 'D', 382-386, 'H', 388-444, 'HN', 447-
A.Cross-references: EMBL:X08030
A.Note: part of this sequence was confirmed by protein sequencing
C:Superfamily: tenascin, EGF homology, fibrinogen beta/gamma homology; fibronectin type
C:Keywords: alternative splicing; calcium binding; cell adhesion; duplication; extracellular
F:1-22/Domain: signal sequence #status predicted <PRO>
F:23-33/Domain: signal sequence #status predicted <PRO>
F:34-181/Product: tenascin 230K #status predicted <MAT>
F:223-249/Domain: EGF homology <EGF1>
F:316-342/Domain: EGF homology <EGF>
F:592-672/Domain: fibronectin type III repeat homology <FN3A>
F:681-763/Domain: fibronectin type III repeat homology <FN3B>
F:773-857/Domain: fibronectin type III repeat homology <FN3C>
F:865-949/Domain: fibronectin type III repeat homology <FN3D>
F:957-1037/Domain: fibronectin type III repeat homology <FN3E>
F:1046-1128/Domain: fibronectin type III repeat homology <FN3F>
F:1137-1219/Domain: fibronectin type III repeat homology <FN3G>
F:1228-1310/Domain: fibronectin type III repeat homology <FN3H>
F:1318-1399/Domain: fibronectin type III repeat homology <FN3I>
F:1407-1487/Domain: fibronectin type III repeat homology <FN3J>
F:1495-1575/Domain: fibronectin type III repeat homology <FN3K>
F:1590-1798/Domain: fibrinogen beta/gamma homology <FBG>
F:1734-1747/Domain: calcium binding #status predicted <CAB>

Query Match 13.4%; Score 151; DB 1; Length 1810;
Best Local Similarity 24.2%; Pred. No. 0.00023;
Matches 61; Conservative 21; Mismatches 78; Indels 92; Gaps 18;

QY 7 POCQAANKTCPTNYMNMNHC---RCLAOEDF-----MFSSDAGDS--TDG----F 49
DB 185 PNC--SEAPCPRNCL-NRGLCVRAKCIPEGFTGEDCSQARPCPDNDGCVGCVCF 241
QY 50 HDICGKNKLEDETC-----OCVCRAGLRPASCG---PKHLEDRNSC---Q 89
DB 242 EGYTGP--DCGELCPHGGCIGRCVGGRCVCHGFTGEDCNEPLCPNCHNRGCVNE 299
QY 90 CVCNKLFPSSOGC---ANRPEDENTC---OCVCR-----TCPRNPLN---P 128
DB 300 CVCDEGYTGEDGELICPDCFRGRCINGTCFCEGYTGEDGELTCNNCNGNRCN 359
QY 129 GKCAC-----ECTES--PQCLKLGKKFHHQTCSCY-----RRP 160
DB 360 GLCVCHGEGVGDGSCQKRCPTKCNNGRCVDGR-CVCHGEGYLGEDGELRCPNDCNHRGR 418
QY 161 CTNRKACEPGF 172
DB 419 CINGCVCDEBGF 430
RESULT 5
A56175
adhesive plaque protein Mgf2 precursor - Mediterranean mussel
C:Species: Mytilus galloprovincialis (Mediterranean mussel)
C>Date: 27-Apr-1995 #sequence_revision 03-Oct-1995 #text_change 11-Jan-2000
C:Accession: A56175
R:Inoue, K.; Takeuchi, Y.; Miki, D.; Odo, S.
J. Biol. Chem. 270, 6698-6701, 1995
A>Title: Mussel adhesive plaque protein gene is a novel member of epidermal growth factor
A:Reference number: A56175; MUID:95204464
A:Accession: A56175
A:Molecule type: mRNA

A:Residues: 1-473 <INO>
A.Cross-references: GB:DA3794; NID:9602767; PIDN:BA07852.1; PID:dl008438; PID:960276
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: duplication
F:1-17/Domain: signal sequence #status predicted <SIG>
F:387-419/Domain: EGF homology <EGF1>
F:429-460/Domain: EGF homology <EGF>
F:23,36,43,56,75,382,424,455,468,473/Modified site: 3',4'-dihydroxyphenylalanine (Tyr

Query Match 13.0%; Score 146.5; DB 2; Length 473;
Best Local Similarity 21.5%; Pred. No. 0.0002;
Matches 67; Conservative 26; Mismatches 83; Indels 135; Gaps 17;

QY 5 TLPOCAANKTCPTNYMNMN-----NHICRL-----AQDFPF---SSDAGD 44
DB 150 TGRRCV--HACKPNCKNKGRCFPGDKTYGRCVDSYSGPTQDNACKPNCSNGRC 207
QY 45 STDFDHD-----ICGPNKLEDETC-----OCVCRAGLRPASCG-- 77
DB 208 SADKREDYSCERPRGYFGPECEHYVCAPNCPKNGIGICSSDGSIGRCRCGYSPTCKY 267
QY 78 ----GPHKELDR---NSCQVCNKLFPSSCGAN-----REPENT----- 111
DB 268 NVCKPFPCKNSGRCVKKSSYNICCKGYSGPTCGENVCKPNPCQNRGCPDNSDGRK 327
QY 112 CCQV-----CK---RTCPRNQPLNPKGC-----ACECT----- 136
DB 328 CRVGVGKPTCECDKPNPNTKPKCKNGKCNKNGKTYTCKCAYGWRGRCTDKAYVNPNC 387
QY 137 -----ESPQKILNKKR-----HQTSCSYR-----RPCTNRKACEPGFY 174
DB 388 VVSKPCKNKGKCIWMNKAIRCKAIVGGRHCKKSKKKNPCASRRCKNRGCTDGNQY 447
QY 175 SEEVCRCPST 185
DB 448 ---VCKCARGY 455

RESULT 6
A49175
Moth B protein - mouse (fragment)
N:Alternate names: Notch homolog
C:Species: Mus musculus (house mouse)
C>Date: 21-Jan-1994 #sequence_revision 05-Jan-1996 #text_change 20-Sep-1999
C:Accession: A49175; P15170; S32113
R:Lairdell, M.; Lendahl, U.
Exp. Cell Res. 204, 364-372, 1993
A>Title: Moth A and Moth B--two mouse Notch homologues coexpressed in a wide variety
A:Reference number: A49175; MUID:93178563
A:Accession: A49175
A>Status: preliminary; nucleic acid sequence not shown
A.Molecule type: mRNA
A.Residues: 1-1203 <LAR>
A.Cross-references: EMBL:X68279; NID:g287989; PIDN:CAA48340.1; PID:g287990
A:Experimental source: embryo
A>Note: sequence extracted from NCBI backbone (NCBI:P126158)
C:Comment: This protein has many EGF repeats and Lin-12/Notch repeats.
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol
F:143-174/Domain: EGF homology <EGX1>
F:482-513/Domain: EGF homology <EGF1>
F:560-591/Domain: EGF homology <EGF>
F:674-705/Domain: EGF homology <EGX2>
F:712-743/Domain: EGF homology <EGF3>
F:836-867/Domain: EGF homology <EGX3>

Query Match 12.3%; Score 139.5; DB 2; Length 1203;
Best Local Similarity 23.8%; Pred. No. 0.0013;
Matches 62; Conservative 16; Mismatches 68; Indels 115; Gaps 15;

QY 16 CPTNYMNMNHCRL-----AQDFPFSSDAGD-----STDG-FHDIC----- 53

F:2009-2041/Domain: ankryrin repeat homology <AN5>

Query Match 12.3% Score 139.5; DB 2; Length 2471;

Best Local Similarity 23.8%; Pred. No. 0.0022; Matches 62; Conservative 16; Mismatches 68; Indels 115; Gaps 15;

16 CPTNNMNNHICRCL-----AQEDFMSSDAGD-----STDG-FPHDIC----- 53
 390 CPTNPLNGYICTCPQATKAGACTEDVDCAAMANSNPCHAKCYNTDGAFFICECLKGA 449
 54 GPNKELD-----EETCQ-----CYCRAGLRPASGPKHELDNRSCQ---CV- 91
 450 GPCREMDINECHSDPCQNDATCLDKIGFTCLCMPEGFK---GVHCELEVNQCQSNPCVN 505
 92 ---CNKKLFSSCGANREDEMTCCQCVCKRTCPRNQPLNPGK-----ACECTE----- 137
 506 NGCQVCKVNRFCQCLPPEGTGVPVCO-IDIDDCSSPCLNGAKCIDHPNGYECQCATGFTG 564
 138 -----SPQKCLLKGKRFPHQTC-----SCYRRPCT 162
 565 TLCDENINDCDPPC-----HHGQCQDGDIDSYTCICNPGYMGALCSDQIDECTISSPCL 617
 163 N-----RQKACEGFS 173
 618 NDGRCTDLVNGYQCMCPETS 638

RESULT 10

A43902
 tenascin - eastern newt (fragment)
 C:Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 12-Feb-1999
 C:Accession: A43902
 R:Onda, H.; Poulin, M.L.; Tassava, R.A.; Chin, I.M.
 Dev. Biol. 148, 219-232, 1991
 A:Title: Characterization of a new tenascin cDNA and localization of tenascin mRNA during development
 A:Reference number: A43902; MUID:92038434
 A:Accession: A43902
 A:Molecule type: mRNA
 A:Residues: 1-647 <OND>
 A:Cross-references: GB:M76615
 A>Note: sequence extracted from NCBI backbone (NCBIN:64543, NCBI:64547)
 C:Superfamily: tenascin; EGF homology; fibronectin beta/gamma homology; fibronectin type C; keywords: extracellular matrix; glycoprotein; tandem repeat
 F:287-313/Domain: EGF homology <EGF>
 F:346-427/Domain: fibronectin type III repeat homology <FN1>
 F:435-517/Domain: fibronectin type III repeat homology <FN2>

Query Match 12.2% Score 138; DB 2; Length 647;

Best Local Similarity 28.2%; Pred. No. 0.0011; Matches 55; Conservative 11; Mismatches 73; Indels 56; Gaps 13;

16 CPTNNMNNHICRCLAQDFMSSDAGDSTDFHICGPNKELDETC---QCVCRAGL 72
 39 CPNNC---NNGRCVDEDCVDEGFTGDDCE---LIC-PNDCEPRGRCINVCVCDCEGF 91
 73 RPASG---PKHELDNRSC---QCVCKNKKLFPSQCG---ANREDEMTC---QCVCKR 118
 92 TGEDEGELTCENNNRRCVAGLCVDDGFOGDDCSSELRCPNDCNDRGRVCNCKVCKE 151
 119 -----TCPRNQPLNPGKCA---CECTESPOKCLLKGKRFPHQSCYRRP----- 160
 152 GPMGEDCADLRCP-MDCNNRGRVCNVCVCDCEG-----FMGEDCSDLRCPCGDCNN 200
 161 ---CTNRKACEGPF 172
 201 RGRVCNVCVCDCEGF 215

RESULT 11
 S42612

transmembrane protein precursor - zebra fish

C:Species: Brachydanio rerio (zebra fish)

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Sep-1999

C:Accession: S42612

R:Bierkamp, C.; Campos-Ortega, J.A.

Mech. Dev. 43, 87-100, 1993

A:Title: A zebrafish homologue of the Drosophila neurogenic gene Notch and its patter

A:Reference number: S42612; MUID:94128602

A:Accession: S42612

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-2437 <BIE>

A:Cross-references: EMBL:X69088; NID:q433866; PIDN:CAA48831.1; PID:q433867

C:Superfamily: unassigned ankryrin repeat proteins; ankryrin repeat homology; EGF homol

F:1023-1054/Domain: EGF homology <EGF1>

F:1185-1216/Domain: EGF homology <EGF2>

F:1915-1947/Domain: ankryrin repeat homology <AN1>

F:1948-1980/Domain: ankryrin repeat homology <AN2>

F:1982-2014/Domain: ankryrin repeat homology <AN3>

F:2015-2047/Domain: ankryrin repeat homology <AN4>

F:2048-2080/Domain: ankryrin repeat homology <AN5>

Query Match 12.2% Score 137.5; DB 2; Length 2437;

Best Local Similarity 22.9%; Pred. No. 0.0031; Matches 57; Conservative 13; Mismatches 66; Indels 113; Gaps 15;

44 DSTGDFHIDICP-----NKELDE-----ETC-----OCVCRAGLRPASC-GP 79
 655 DKINGEYCECEPYSGSMCNINIDICALNPGCHNGGTCIDGVNSFTCLCPDGFREDATCLISQ 714
 80 HKELDNRSC-----OCVCKNKKLFPSQCGAN-REPELNTC-----Q 113
 715 HNEGSSNPCTHSGCILDQINSYRCVCEAGMGRRNDININECLSNPCVNGCTCKMTSGYL 774
 114 CVCK-----RNCPRNQPLNPGKC-----ACECTE-----SPQ 140
 775 CTCRAGFSGPNQMNINBASNPCLNGSCIDVAGFKNCMLPTGVECVNLAPCSFR 834
 141 KCLKG-----KRFPHQTC-----SCYRRPCTNRQACE---PQFSYSE 176
 835 PCKNGCVRESEDFOSFCNCPAGNQGTCVEDINECVRNPTN-GVCEINLRGGFO--- 890
 177 EYCRGVPSY 185
 891 --CRCNPGP 897

RESULT 12

A35844

Xotch protein - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C:Date: 12-Oct-1990 #sequence_revision 12-Oct-1990 #text_change 13-Aug-1999

C:Accession: A35844

R:Coffman, C.; Harris, W.; Kintner, C.

Science 249, 1438-1441, 1990

A:Title: Xotch, the Xenopus homologue of Drosophila notch.

A:Reference number: A35844; MUID:90385285

A:Accession: A35844

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

A:Molecule type: mRNA

A:Residues: 1-2524 <COF>

C:Superfamily: unassigned ankryrin repeat proteins; ankryrin repeat homology; EGF homol

C:Keywords: transmembrane protein

F:146-177/Domain: EGF homology <EGX1>

F:184-215/Domain: EGF homology <EGF1>

F:222-254/Domain: EGF homology <EGF2>

F:456-487/Domain: EGF homology <EGF3>

F:757-788/Domain: EGF homology <EGF4>

F:1025-1056/Domain: EGF homology <EGF5>

F:1924-1956/Domain: ankryrin repeat homology <AN1>

F:1957-1989/Domain: ankryrin repeat homology <AN2>

F:1991-2023/Domain: ankyrin repeat homology <AN3>
F:2024-2056/Domain: ankyrin repeat homology <AN4>
F:2057-2089/Domain: ankyrin repeat homology <AN5>

Query Match 12.2%; Score 137.5; DB 2; Length 2524;
Best Local Similarity 21.7%; Pred. No. 0.0031;
Matches 55; Conservative 21; Mismatches 77; Indels 101; Gaps 12;

OY 7 POCQANKTCPTNYMNN-----HICRL-----AQEDFMFSSDAGDDSTGPHDI 52
DB 766 PNCQNNINECSSNPLNHTCIDVAGKCNMPLYTGAIICAVLAPCAGSPCKNG----- 841
OY 53 CGPNKELDE-ETFCOCVCRAGLRPASCGPHKELDRNSC-----OCVCK 93
DB 842 -GRCESEDEFETFCSCPPGMOGOTC-----EIDMNECVNRPCRNATCONTNGSKCNCK 896
OY 94 NKLPFSGCAGNREFDENTCOCVCKRTCPRNQPLNPGKCA-----CCTESPOKCLLKG 146
DB 897 ---PGYTGRCNEMDIDCQ-----PNCCHNGSGSCSDGINMFECNCPAG-----FRG 939
OY 147 KFFHHQTCSCYRPCTNRCK-----ACEPGFS----- 173
DB 940 PCCEEDINECASNPKNGANCTDCVNSYTCQPGFSGIHCESNTPDCTESSCFNGGTCTI 999
OY 174 --YSEVCRCPVPSY 185
DB 1000 DGINTFTCCCPGPF 1013

RESULT 13

A61625
tenascin-like protein precursor - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 21-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 11-Jan-2000
C:Accession: A61625; S28463
R:Baumgartner, S.; Chiquet-Ehrismann, R.
Mech. Dev. 40, 165-176, 1993
A:Title: Ten(a), a Drosophila gene related to tenascin, shows selective transcript local
A:Reference number: A61625; MUID:93264270
A:Accession: A61625
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-782 <BAU>
A:Cross-references: EMBL:X68794
C:Genetics:
A:Gene: ten-a
A:Cross-references: FlyBase:FBgn0004446
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: tandem repeat
F:1-62/Domain: signal sequence #status predicted <SIG>
F:63-782/Product: tenascin-like protein #status predicted <MAN>
F:497-524/Domain: EGF homology <EGF>

Query Match 12.0%; Score 136; DB 2; Length 782;
Best Local Similarity 23.4%; Pred. No. 0.0018;
Matches 52; Conservative 19; Mismatches 75; Indels 76; Gaps 13;

OY 6 LPQCAANK-----TCPTNYMNNHICRCLAQDFMFSSDAGDSTDGPHDTCGPKKELDE 61
DB 499 VPCNSHRCIEGECHECRGMGPYC-----DQHDCLDPLCS-----GH 537
OY 62 ETC---QCVRAGLRPASCGPHKELDRNSCOCVCKNKLFPSSQGANREFDENTCOCVCKR 118
DB 538 GTCVAGQCYCKRAGMGEDCG---TIDQOYQCL-----PCSHGTYTDLTGQCVCKR 587
OY 119 -----TCPRNQLNPGKCA-----CTESP--OKCLLKGKFFHHQCS 155
DB 588 HMTGPDSCQAVCSLDCGRNGVCESGKCRNSGWTGNLDCQLPCDSRCEHG--QCKNGTCV 646
OY 156 CYR-----RPTNRKQACRPGFS-----YSEVCRCPVPSY 185

DB 647 CSQGNNGRHCT--LPCENGCSRHGQCTLENGEYRCDCIEGW 686

RESULT 14

A57278
fibrillin-2 precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 11-Jan-2000
C:Accession: A57278
R:Zhang, H.; Hu, W.; Ramirez, F.
J. Cell Biol. 129, 1165-1176, 1995
A:Title: Developmental expression of fibrillin genes suggests heterogeneity of extrac
A:Reference number: A57278; MUID:95263670
A:Accession: A57278
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2907 <ZHA>
A:Cross-references: GB:L39790; NID:g762830; PIDN:AAA74908.1; PID:g762831
C:Superfamily: unassigned EGF-related proteins; EGF homology
F:1239-1274/Domain: EGF homology <EGF1>
F:2488-2523/Domain: EGF homology <EGF>

Query Match 12.0%; Score 136; DB 2; Length 2907;
Best Local Similarity 23.9%; Pred. No. 0.0045;
Matches 60; Conservative 24; Mismatches 83; Indels 84; Gaps 15;

OY 6 LPQCAANKTCPTNYMNN-----HICRC-----LQEDFMFSSDAGDDST----- 46
DB 2485 IDECSQSPKPC--NFCIKRTKGSYQSCPRGYVLOEDKRTCKDLDECOTKOHNCQFLCVN 2542
OY 47 -----DGF---HDICGPKELDEE-----TCQ-----CVRAGL-----RP 74
DB 2543 TLGFTCKCPREGFTQHNHACIDNNECGSQLCAKGLICQNTPGSFSCQCRFSLDASG 2602
OY 75 ASCGPHKELDRN-SCQCVCKNKLFPSSQGA-----NREFDENTCOCVCKRTCPRN 123
DB 2603 LNCEVDDEDCDGNHRCQHCQNTLGGYRCGCPHGDVQHYQWNCVDNEDCS----- 2652
OY 124 QPLNPGK---ACECTESPOKCLL-KGKFFHHQTCSCYR-RPCTNRKQACRPGFSYSE-- 176
DB 2653 ---NPGACGSASCTYTLGSLYKQACPSGFSFDQFSACHDVNEDCSSKNPCSTGCSNTEGG 2709
OY 177 EYVCRCPVPSYWK 187
DB 2710 YLCCGCPGYFR 2720

RESULT 15

T18355
hypothetical protein P3 - Mycoplasma hyorhinis
C:Species: Mycoplasma hyorhinis
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 07-Dec-1999
C:Accession: T18355
R:Deng, G.; McIntosh, M.A.
J. Bacteriol. 176, 5929-5937, 1994
A:Title: An amplifiable DNA region from the Mycoplasma hyorhinis genome.
A:Reference number: Z18888; MUID:95014025
A:Accession: T18355
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1187 <DEN>
A:Cross-references: EMBL:L11447; NID:g150156; PID:g664956; PIDN:AAA62228.1
C:Genetics:
A:Genetic code: SGC3

Query Match 12.0%; Score 135.5; DB 2; Length 1187;
Best Local Similarity 24.9%; Pred. No. 0.0027;
Matches 47; Conservative 17; Mismatches 48; Indels 77; Gaps 13;

OY 26 TORCLAQEDFMFSSDAG-----DSTDGPH-----DICGPKLEDEFTCC- 66

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:51:43 ; Search time 42.92 Seconds

(without alignments)
153.240 Million cell updates/sec

Title: US-09-427-657-2_COPY_228_419
Perfect score: 1130
Sequence: 1 SLPTATLPCQCAANKTCPTNY.....SYSEVCGCVSYWKRPQMS 192

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1330	100.0	419	1	VEGC_HUMAN
2	998	88.3	415	1	VEGC_MOUSE
3	238	21.1	1700	1	BAR3_CHITE
4	148	13.1	1808	1	TENA_CHICK
5	146.5	13.0	473	1	FP2_MYTGA
6	140.5	12.4	2813	1	VWF_CANFA
7	139.5	12.3	1746	1	TENA_PIG
8	137.5	12.2	2437	1	NOTC_BRARE
9	136	12.0	2524	1	NOTC_XENLA
10	135	11.9	2907	1	FN2_MOUSE
11	135	11.9	2911	1	FN2_HUMAN
12	134.5	11.9	2482	1	VWF_PIG
13	130	11.5	931	1	EMR1_MOUSE
14	130	11.5	2201	1	TENA_HUMAN
15	128.5	11.4	1696	1	PKC5_BRACL
16	128.5	11.4	2871	1	FN1_HUMAN
17	128.5	11.4	2871	1	FN1_MOUSE
18	128	11.3	2703	1	NOTC_DROME
19	127	11.2	2871	1	FN1_BOVIN
20	126	11.2	3635	1	LMA5_MOUSE
21	125	11.1	2444	1	NTC1_HUMAN
22	124	11.0	570	1	FBP3_STRPU
23	124	11.0	3712	1	LMA_DROME
24	123.5	10.9	220	1	AMTA_HYDRA
25	123.5	10.9	5179	1	MYC2_HYDRA
26	123	10.9	1106	1	STC_DROME
27	123	10.9	2813	1	NTC3_MOUSE
28	123	10.9	2813	1	VWF_HUMAN
29	122.5	10.8	1877	1	PKC5_MOUSE
30	122.5	10.8	2531	1	NTC1_MOUSE
31	121	10.7	1680	1	FUR2_DROME
32	120.5	10.7	769	1	ITB2_PIG
33	120.5	10.7	2139	1	CRB_DROME

ALIGNMENTS

RESULT	ID	VEGC_HUMAN	STANDARD	PRT	419 AA.
34	120	10.6	787	1	ITB3_MOUSE
35	119	10.5	2531	1	NTC1_RAT
36	118	10.4	1104	1	NEPL_HUMAN
37	118	10.4	3672	1	LMI2_CAEEL
38	118	10.4	4289	1	TENX_HUMAN
39	116	10.3	1964	1	NTC4_MOUSE
40	115.5	10.2	1394	1	TGFB_HUMAN
41	115	10.2	379	1	WIF1_HUMAN
42	115	10.2	4544	1	LRL1_HUMAN
43	114.5	10.1	1403	1	NID2_MOUSE
44	114	10.1	326	1	BAR1_CHIPA
45	114	10.1	1064	1	FBP1_STRPU

054890 mus musculus
007008 rattus norv
012986 homo sapien
021313 caenorabd1
P22105 homo sapien
P31695 mus musculu
Q955W5 homo sapien
Q07954 homo sapien
O88322 mus musculu
P08724 chironomus
P10079 strongyloe

RESULT 1
VEGC_HUMAN
ID VEGC_HUMAN STANDARD: PRT: 419 AA.
AC P49767: 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR (VEGF-C) (VASCULAR
DE ENDOTHELIAL GROWTH FACTOR RELATED PROTEIN) (VRE) (FLT4 LIGAND) (FLT4-
DE 1).
GN VEGFC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 103-120.
RX MEDLINE=96178224; PubMed=8617204;
RA Joukov V., Pajusola K., Kaipainen A., Chilov D., Lahtinen I., Kukk E.,
RA Saksela O., Kalkkinen N., Alitalo K.;
RT "A novel vascular endothelial growth factor, VEGF-C, is a ligand for
RT the FLT4 (VEGFR-3) and KDR (VEGFR-2) receptor tyrosine kinases.";
RL EMBO J. 15:1751-1751(1996).
[2]
RN ERRATUM.
RP MEDLINE=96203094; PubMed=8612600;
RX Joukov V., Pajusola K., Kaipainen A., Chilov D., Lahtinen I., Kukk E.,
RA Saksela O., Kalkkinen N., Alitalo K.;
RL EMBO J. 15:1751-1751(1996).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96312526; PubMed=8700872;
RA Lee J., Gray A., Yuan J., Luoh S.-M., Avraham H., Wood W.I.;
RT "Vascular endothelial growth factor-related protein: a ligand and
RT specific activator of the tyrosine kinase receptor FLT4.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:1988-1992(1996).
[4]
RP SEQUENCE FROM N.A.
RA Fitz L., Morris J.C., Towler P.S., Long A.J., Greco R.,
RA Burgess P., Giampotti J., Charette A., Hennessey D., Kovacic S.,
RA Fitzgerald M., Scaltreco H., Welch N., Neben S., Finnelly H.,
RA Zollner R., Wang J., Nickbarg E., Gassaway R., Turner K.,
RA Wood C.R.;
RT Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
CC CELL GROWTH.
CC - SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC - PTM: PROBABLY PROTEOLITICALLY PROCESSED IN THE C-TERMINUS.
CC - SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@lsb-sib.ch).

CC EMBL: X94216; CAA63907.1; -
 CC EMBL: U43142; AAB85214.1; -
 DR EMBL: U58111; AAB02909.1; -
 DR HSP: P15692; 1VPE.
 DR MIM: 601528; -
 DR InterPro: IPR000072; -
 DR InterPro: IPR002400; -
 DR Pfam: PF00341; PDGF_1;
 DR PROSITE: PS00249; PDGF_1;
 DR PROSITE: PS00278; PDGF_2; 1.
 DR Mitogen; Growth factor; Glycoprotein; Signal; Repeat.
 KW SIGNAL
 FT PROPEP 1 102 POTENTIAL.
 FT CHAIN 103 419 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
 FT DOMAIN 275 365 4 X 24 AA TANDEM REPEATS.
 FT REPEAT 275 298 1.
 FT REPEAT 299 322 2.
 FT REPEAT 323 346 3.
 FT REPEAT 347 365 4 (PARTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 419 AA; 46883 MW; 9F598719DB3E014F CRC64;

Query Match 100.0%; Score 1130; DB 1; Length 419;
 Best Local Similarity 100.0%; Pred. No. 1.7e-87;
 Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLPTATLPQOQAAANKTCPTNYMNMNHHICRLAODEMFSSDAGDDSTGDFHDCGPNKELD 60
 DB 228 SLPTATLPQOQAAANKTCPTNYMNMNHHICRLAODEMFSSDAGDDSTGDFHDCGPNKELD 287
 QY 61 EETCCGVCAGLRPASCGRHKELDNSCQVCNKLFPSCGAGNREFDENTCQVCCKRTC 120
 DB 288 EETCCGVCAGLRPASCGRHKELDNSCQVCNKLFPSCGAGNREFDENTCQVCCKRTC 347
 QY 121 PNNOPNPGKACCECTESPQCKLKGKRFHQTSCYRRPCTNRKACBPGEFSYSSEVCR 180
 DB 348 PNNOPNPGKACCECTESPQCKLKGKRFHQTSCYRRPCTNRKACBPGEFSYSSEVCR 407
 QY 181 CVPSTYKRRPQMS 192
 DB 408 CVPSTYKRRPQMS 419

RESULT 2
 VEGF_MOUSE STANDARD; PRT; 415 AA.
 AC P97953;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR (VEGF-C) (FLT4 LIGAND)
 DE (FLT4-L).
 GN VEGFC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BAIB/C;
 RX MEDLINE=97164697; PubMed=9012504;
 RA Kulk E., Lymboussaki A., Taira S., Kaipainen A., Jeltsch M.,
 RA Joukov V., Altai K.;
 RT "VEGF-C receptor binding and pattern of expression with VEGFR-3
 RT suggests a role in lymphatic vascular development.";
 RL Development 122:3829-3837(1996).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=BAIB/C;
 RX MEDLINE=97388482; PubMed=9247316;
 RA Fitz J.J., Morris J.C., Towler P., Long A., Burgess P., Greco R.,
 RA Wang J., Gassaway R., Nickbarg E., Kovacic S., Ciavetta A.,
 RA Giannotti J., Finerty H., Zollner R., Belier D.R., Leek L.V.,
 RA Turner K.J., Wood C.R.;
 RT "Characterization of murine Flt4 ligand/VEGF-C.";
 RL Oncogene 15:613-618(1997).
 CC -1- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
 CC CELL GROWTH.
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).

DR EMBL: U73620; AAC52984.1; -
 DR EMBL: U58112; AAB46707.1; -
 DR HSP: P15692; 1VPE.
 DR MGD: MGI:109124; Vegfc.
 DR InterPro: IPR000072; -
 DR InterPro: IPR002400; -
 DR Pfam: PF00341; PDGF_1;
 DR PRINTS: PR00438; GFCYSKNOT.
 DR PROSITE: PS00249; PDGF_1;
 DR PROSITE: PS00278; PDGF_2; 1.
 DR Mitogen; Growth factor; Glycoprotein; Signal; Repeat.
 KW SIGNAL
 FT PROPEP 1 98 POTENTIAL.
 FT CHAIN 99 415 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
 FT DOMAIN 271 361 4 X 24 AA TANDEM REPEATS.
 FT REPEAT 271 294 1.
 FT REPEAT 295 318 2.
 FT REPEAT 319 342 3.
 FT REPEAT 343 361 4 (PARTIAL).
 FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 415 AA; 46471 MW; D9D3DD3C8CC659D6 CRC64;

Query Match 88.3%; Score 998; DB 1; Length 415;
 Best Local Similarity 84.4%; Pred. No. 1.6e-76;
 Matches 162; Conservative 18; Mismatches 12; Indels 0; Gaps 0;

QY 1 SLPTATLPQOQAAANKTCPTNYMNMNHHICRLAODEMFSSDAGDDSTGDFHDCGPNKELD 60
 DB 224 SLPTATLPQOQAAANKTCPTNYMNMNHHICRLAODEMFSSDAGDDSTGDFHDCGPNKELD 283
 QY 61 EETCCGVCAGLRPASCGRHKELDNSCQVCNKLFPSCGAGNREFDENTCQVCCKRTC 120
 DB 284 EETCCGVCAGLRPASCGRHKELDNSCQVCNKLFPSCGAGNREFDENTCQVCCKRTC 343
 QY 121 PNNOPNPGKACCECTESPQCKLKGKRFHQTSCYRRPCTNRKACBPGEFSYSSEVCR 180
 DB 344 PNNOPNPGKACCECTESPQCKLKGKRFHQTSCYRRPCTNRKACBPGEFSYSSEVCR 403
 QY 181 CVPSTYKRRPQMS 192
 DB 404 CVPSTYKRRPQMS 415

RESULT 3
 BAR3_CHITE STANDARD; PRT; 1700 AA.
 ID BAR3_CHITE
 AC Q03376;
 DT 01-OCT-1993 (Rel. 27, Created)

```

DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE BALBIANI RING PROTEIN 3 PRECURSOR.
GN BR3.
OS Chironomus tentans (Midge).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomidae; Chironomidae; Chironominae; Chironomus.
OX NCBI_TaxID=7153;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=90172404; PubMed=1689777;
RA Paulsson G., Lendahl U., Galli J., Ericsson C., Mieslander L.;
RT "The Balbiani ring 3 gene in Chironomus tentans has a diverged
RT repetitive structure split by many introns.";
RL J. Mol. Biol. 211:331-349(1990).
CC -1- FUNCTION: USED BY THE LARVAE TO CONSTRUCT A SUPRAMOLECULAR
CC STRUCTURE, THE LARVAL TUBE. BALBIANI RING PROTEIN 3 COULD PLAY A
CC ROLE AS A TRANSPORT PROTEIN THAT BINDS TO OTHER PROTEINS
CC INTRACELLULARLY AND IN THE GLAND LUMEN IN ORDER TO PREVENT THESE
CC FROM FORMING WATER-INSOLUBLE FIBERS TOO EARLY.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: SALIVARY GLAND.
CC -1- DOMAIN: HAS 82 APPROXIMATE REPEATS OF CYS-X-CYS-X-CYS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X52263; CA36506.1; -
DR PIR: S08167; S08167.
DR HSSP: P18055; 2MRB.
DR InterPro: IPR000853; -
DR PRINTS: PR00876; MTNEMATODE.
DR KM Repeat: Signal.
FT SIGNAL. 1 2
FT CHAIN. 1 1700 POTENTIAL.
FT SEQUENCE 1700 AA; 186145 MW; 342028521B0815 CRC64;
SQ

```

Query Match 21.1%; Score 238; DB 1; Length 1700;
 Best Local Similarity 25.4%; Pred. No. 1e-12; Indels 82; Gaps 9;
 Matches 62; Conservative 26; Mismatches 74;

```

OY 3 PATLPQCCQANK-----TCPTNYMNNHICRCLAQEDPFMSDAGDGS 45
DB 1204 PTPAPPCSNMOKYNSVSCGCGNPKRGKCGNQIWCNTRCVCPRKMEKPRADN---- 1259
OY 46 TPGFHIDICPKNELDEBTQCVCYCRAG-----LRPASCGPH 80
DB 1260 -----C-KTVMWMDQOCQVCPCGEGCGKGYKMNANTCSECPADKAPASCGDK 1311
OY 81 KELDRNSOCQVCKNLFPSOGCANREEDENTGOCYCKRT--CPRNPPLN-----P 128
DB 1312 KSWNDSCSCQCKSKMPCCGCPNNOQMNKDECKCSATGNCPAQQVMSQTCGSCSPT 1371
OY 129 GKC-----ACECTESPCKLKGKKFHQTCSCY---RPCTNRKAKCEGFGFSY 175
DB 1372 GKCCTAGQVWCSSACACVCYPAQKKCDSPKTDWNSCSCQCKPKMRPPTG---CGNAGRITWD 1428
OY 176 EEVC 179
DB 1429 DATC 1432

```

RESULT 4
 TENA_CHICK STANDARD: PRT; 1808 AA.

```

AC P10039; P13132; O73584; O73585;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TENASCIN PRECURSOR (TN) (HEXABRACHION) (CYTOTACTIN) (NEURONECTIN)
DE (GEMM) (JI) (MIOTENDINOUS ANTIGEN) (GILOMA-ASSOCIATED-EXTRACELLULAR
DE MATRIX ANTIGEN) (GP 150-225).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=90030407; PubMed=2478295;
RA Spring J., Beck K., Chiquet-Ehrismann R.;
RT "Two contrary functions of tenascin: dissection of the active sites
RT by recombinant tenascin fragments.";
RL Cell 59:325-334(1989).
RN [2]
RP SEQUENCE OF 27-722 FROM N.A., AND SEQUENCE OF 79-96.
RC TISSUE=Fibroblast;
RX MEDLINE=89030589; PubMed=2460335;
RA Pearson C.A., Pearson D., Shibahara S., Hofsteenge J.,
RA Chiquet-Ehrismann R.;
RT "Tenascin: cDNA cloning and induction by TGF-beta.";
RL EMBL J. 7:2977-2982(1986).
RN [3]
RP SEQUENCE OF 464-1018 AND 1412-1661 FROM N.A., AND SEQUENCE OF 852-868.
RC TISSUE=Embryo;
RX MEDLINE=88176910; PubMed=2451243;
RA Jones F.S., Burgeon M.P., Hoffman S., Crossin K.L., Cunningham B.A.,
RA Edelman G.M.;
RT "A cDNA clone for cytotactin contains sequences similar to epidermal
RT growth factor-like repeats and segments of fibronectin and
RT fibronogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:2186-2190(1988).
CC -1- FUNCTION: SAM (SUBSTRATE-ADHESION MOLECULE) THAT APPEARS TO
CC INHIBIT CELL MIGRATION. MAY PLAY A ROLE IN SUPPORTING THE GROWTH
CC OF EPITHELIAL TUMORS.
CC -1- SUBUNIT: HEXAMERIC. A HOMOTRIMER MAY BE FORMED IN THE TRIPLE
CC COILED-COIL REGION AND MAY BE STABILIZED BY DISULFIDE RINGS AT
CC BOTH ENDS. TWO OF SUCH HALF-HEXABRACHIONS MAY BE DISULFIDE LINKED
CC WITHIN THE CENTRAL GLOBULE.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 230 KDA (SHOWN HERE), 200 KDA
CC AND 190 KDA; ARE PRODUCED BY ALTERNATIVE SPLICING. THEY ARE
CC PRODUCED IN A TISSUE-AND TIME-SPECIFIC MANNER DURING DEVELOPMENT.
CC -1- INDUCTION: BY TGF-BETA.
CC -1- SIMILARITY: CONTAINS 13.5 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 11 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS A DOMAIN RELATED TO THE C-TERMINAL PART OF
CC THE BETA AND GAMMA CHAINS OF FIBRINOGEN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M23121; AAA49086.1; -
DR EMBL: X08031; CAB40811.1; -
DR EMBL: X08030; CA30824.1; ALT_TERM.
DR EMBL: J03641; AAA48748.1; ALT_SEQ.
DR EMBL: M20816; AAA48749.1; ALT_SEQ.
DR PIR: A30903; A30903.
DR PIR: A31930; A31930.
DR PIR: A33379; A33379.
DR PIR: B33379; B33379.
DR PIR: C33379; C33379.

```

FT	DISULFID	502	512	BY SIMILARITY.
FT	DISULFID	506	517	BY SIMILARITY.
FT	DISULFID	519	528	BY SIMILARITY.
FT	DISULFID	533	543	BY SIMILARITY.
FT	DISULFID	537	548	BY SIMILARITY.
FT	DISULFID	550	559	BY SIMILARITY.
FT	DISULFID	564	574	BY SIMILARITY.
FT	DISULFID	568	579	BY SIMILARITY.
FT	DISULFID	581	590	BY SIMILARITY.
FT	CARBOHYD	38	38	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	168	168	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	328	328	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	603	603	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	643	643	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	751	751	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	759	759	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1050	1050	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1090	1090	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1101	1101	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1112	1112	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1153	1153	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1183	1183	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1416	1416	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1736	1736	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1769	1769	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VANSPLIC	1043	1224	MISSING (IN 200 KDA ISOFORM).
FT	VANSPLIC	1043	1315	MISSING (IN 190 KDA ISOFORM).
FT	CONFLICT	182	182	W -> R (IN REF. 2).
FT	CONFLICT	563	571	SCPDNCNV -> PAPMTATW (IN REF. 3).
FT	CONFLICT	598	598	E -> G (IN REF. 3)
FT	CONFLICT	838	838	T -> TXY (IN REF. 3).
FT	CONFLICT	886	886	N -> Y (IN REF. 3).
QO	SEQUENCE	1808 AA:	198858 MW:	B924A06CF9EFD6DE CRC64;

Query March	13.1%	Score 148:	DB 1:	Length 1808:
Best Local Similarity	24.2%	Pred. No. 3.3e-05:		
Matches 61:	Conservative 20:	Mismatches 79:	Indels 92:	Gaps 18:

QY	7	PCCAAKNTCTPTMMNNHIC	----	RCLAEQF	-----	MSSDAGDS	--	TTG	----	F	49
Db	185	PNC--SEPAECRNCL	-	NGSLCVRGKCI	CEEFTEGEDCSQAC	PSDCNDQGC	VC	GV	CVCF		241
QY	50	HDICGPNKEDEBETC	-----	QCVCRAGLRPASG	----	PKELDRNSC	----	Q	89		
Db	242	EYGTGP--DCGELICPHGCG	ILRGVGRGVC	HGGFTGEDCNE	PLCPNCHNRGR	CV	YDNE		299		
QY	90	CVCKNKLFPSQCG	----	ANREFDENTC	----	QVCYKR	-----	TCRQPLN	----	P	128
Db	300	CVCEBGYTGDECGELIC	PNDCFDRGR	CTNGTCFCE	BEGYTGDECGELT	C	PNNCNG	NGRCEN		359	
QY	129	GKGCAC	-----	ECTES	--	POKCLLKGRKFFHQCSCY	-----	RRP	160		
Db	360	GLCYCHGEFVGDDCSQK	KCRKPCDNNRGH	CVGR	-	CVCHGYTGDECGEL	CRPNDCHNRGR		418		
QY	161	CTNRKACEPGF	172								
Db	419	CINGCVCDEGF	430								

RESULT	5
FP2_MYTGA	
Db	FP2_MYTGA
AC	Q25464:
DT	30-MAY-2000 (Rel. 39, Created)
DT	30-MAY-2000 (Rel. 39, Last sequence update)
DT	01-OCT-2000 (Rel. 40, Last annotation update)
DE	ADHESIVE PLAQUE MATRIX PROTEIN 2 PRECURSOR (FOOT PROTEIN 2) (MGFP2) (MGFP-2).
GN	FP2.
OS	Mytilus galloprovincialis (Mediterranean mussel).
OC	Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;

[illegible]

RT factor and identification of a mutation causing Scottish terrier von
Willebrand's disease.";
RT Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE OF 1234-1669 FROM N.A.
RC TISSUE-Blood;
RA Mancuso D.J., Christopherson P.A., Kroner P.A., Montgomery R.R.;
RT "The canine von Willebrand factor gene: sequence and expression of
a region encoding the glycoprotein Ib/IX binding domain.";
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION: IMPORTANT IN THE MAINTENANCE OF HOMEOSTASIS. IT
PARTICIPATES IN PLATELET-VESSEL WALL INTERACTIONS BY FORMING A
NONCOVALENT COMPLEX WITH COAGULATION FACTOR VIII AT THE SITE OF
VASCULAR INJURY (BY SIMILARITY).
CC - SUBUNIT: MULTIMERIC (BY SIMILARITY).
CC - TISSUE SPECIFICITY: BLOOD.
CC - PTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR
INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).
CC - SIMILARITY: CONTAINS 3 WFMA DOMAINS.
CC - SIMILARITY: CONTAINS 3 WFC DOMAINS.
CC - SIMILARITY: CONTAINS 4 WFMD DOMAINS.
CC - SIMILARITY: CONTAINS A C-TERMINAL CYSTEINE KNOT-LIKE DOMAIN (CTCK).
CC - SIMILARITY: SOME, TO SILKWORM HEMOCYTIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC
DR EMBL: L76227; AAB05549.1; -
DR EMBL: L16903; AAA05903.1; -
DR EMBL: AF099154; AAD04919.1; -
DR EMBL: U66246; AAB93766.1; -
DR HSP: P04275; IATZ.
DR InterPro: IPR000359; -
DR InterPro: IPR001007; -
DR InterPro: IPR001845; -
DR InterPro: IPR002035; -
DR InterPro: IPR002919; -
DR Pfam: PF00007; Cys_Knotc; 1.
DR Pfam: PF01826; TIL; 4.
DR Pfam: PF00092; vwa; 3.
DR Pfam: PF00093; vwc; 3.
DR Pfam: PF00094; vwd; 4.
DR PRINTS: PR00453; VWFADOMAIN.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS50234; WFEA; 3.
DR PROSITE: PS01208; WVEC; 3.
KW Blood coagulation; Platelet; Glycoprotein; Extracellular matrix;
KW Plasma; Endothelial cell; Repeat; Cell adhesion; Signal.
KW SIGNAL.
FT SIGNAL 1 22
FT PROPEP 23 763
FT CHAIN 764 2813
FT DOMAIN 35 179
FT DOMAIN 388 541
FT DOMAIN 764 787
FT DOMAIN 788 833
FT DOMAIN 826 853
FT DOMAIN 867 1013
FT DOMAIN 1277 1453
FT DOMAIN 1498 1665
FT DOMAIN 1691 1871
FT DOMAIN 1950 2102
FT DOMAIN 2216 2261
FT DOMAIN 2255 2326
FT DOMAIN 2429 2495
FT DOMAIN 2580 2650
FT DOMAIN 2724 2812
FT SITE 531 533
CELL ATTACHMENT SITE (POTENTIAL).

FT	SITE	698	700	CELL ATTACHMENT SITE (POTENTIAL).
FT	SITE	2507	2509	CELL ATTACHMENT SITE (POTENTIAL).
FT	DISULFID	767	808	BY SIMILARITY.
FT	DISULFID	776	804	BY SIMILARITY.
FT	DISULFID	810	821	BY SIMILARITY.
FT	DISULFID	867	996	BY SIMILARITY.
FT	DISULFID	889	1031	BY SIMILARITY.
FT	DISULFID	898	993	BY SIMILARITY.
FT	DISULFID	914	921	BY SIMILARITY.
FT	DISULFID	1060	1084	BY SIMILARITY.
FT	DISULFID	1071	1111	BY SIMILARITY.
FT	DISULFID	1089	1091	BY SIMILARITY.
FT	DISULFID	1153	1165	BY SIMILARITY.
FT	DISULFID	1149	1169	BY SIMILARITY.
FT	DISULFID	1126	1130	BY SIMILARITY.
FT	DISULFID	1156	1199	BY SIMILARITY.
FT	DISULFID	1234	1237	BY SIMILARITY.
FT	DISULFID	1272	1458	BY SIMILARITY.
FT	DISULFID	1669	1670	BY SIMILARITY.
FT	DISULFID	1686	1872	BY SIMILARITY.
FT	DISULFID	1879	1904	BY SIMILARITY.
FT	DISULFID	1899	1940	OR 1942 (BY SIMILARITY).
FT	DISULFID	1972	2123	BY SIMILARITY.
FT	DISULFID	1950	2085	BY SIMILARITY.
FT	DISULFID	1927	2088	BY SIMILARITY.
FT	DISULFID	1993	2001	BY SIMILARITY.
FT	DISULFID	2724	2774	BY SIMILARITY.
FT	DISULFID	2739	2788	BY SIMILARITY.
FT	DISULFID	2750	2804	BY SIMILARITY.
FT	DISULFID	2754	2806	BY SIMILARITY.
FT	DISULFID	?	2811	BY SIMILARITY.
FT	CARBOHYD	99	99	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	156	156	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	211	211	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	666	666	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	857	857	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1231	1231	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1515	1515	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1574	1574	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2223	2223	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2290	2290	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2357	2357	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2400	2400	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2546	2546	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2585	2585	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2790	2790	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	55	55	G -> E (IN REF. 2).
FT	CONFLICT	70	70	V -> I (IN REF. 3).
FT	CONFLICT	266	266	A -> G (IN REF. 2).
FT	CONFLICT	280	280	I -> V (IN REF. 2).
FT	CONFLICT	409	411	VCH -> ICO (IN REF. 2).
FT	CONFLICT	994	994	G -> A (IN REF. 1).
FT	CONFLICT	1021	1021	F -> L (IN REF. 2).
FT	CONFLICT	2381	2381	L -> P (IN REF. 2).
FT	CONFLICT	2406	2406	P -> L (IN REF. 2).
SO	SEQUENCE	2813	AA; 309716	MM; 5DF93ELE5E72F80C CRC64;

Query Match 12.4%; Score 140.5; DB 1; Length 2813;
Best Local Similarity 20.5%; Pred. No. 0.0002;
Matches 67; Conservative 37; Mismatches 56; Indels 167; Gaps 21;
QY 15 TCPTNYMNN--HICRCLAQEDFMFSDAGDSTGDFHICGPNKRLDEPTC-----Q 65
DB 2202 SCPPSLVYVHHCHEGPRICEGN---TSQGDQPSSEGC--FCPPNVMLESGVPEACTQ 2256
QY 66 CYCRAGLR-----PA-----SCGPRK--ELDRNS 87
DB 2257 CISEDGVVRHOFLEWTWPAHQPCOICTCLGRKVNCTLOPCPTARAPTCGCEVARLRONA 2316
QY 88 CQCV-----CKNKL-----FSCGAGAREPDENCCG---VCR-- 118
DB 2317 EQCCPEYECVCDLVSCDLPVPVPCEDGLQMTLTNDECRPN---FTVCARKEKDCRRS 2371


```

FT DISULFID 439 449 BY SIMILARITY.
FT DISULFID 443 454 BY SIMILARITY.
FT DISULFID 456 465 BY SIMILARITY.
FT DISULFID 470 480 BY SIMILARITY.
FT DISULFID 474 485 BY SIMILARITY.
FT DISULFID 487 496 BY SIMILARITY.
FT DISULFID 501 511 BY SIMILARITY.
FT DISULFID 505 516 BY SIMILARITY.
FT DISULFID 518 527 BY SIMILARITY.
FT DISULFID 532 542 BY SIMILARITY.
FT DISULFID 536 547 BY SIMILARITY.
FT DISULFID 549 558 BY SIMILARITY.
FT DISULFID 563 573 BY SIMILARITY.
FT DISULFID 567 578 BY SIMILARITY.
FT DISULFID 580 589 BY SIMILARITY.
FT DISULFID 594 604 BY SIMILARITY.
FT DISULFID 598 609 BY SIMILARITY.
FT DISULFID 611 620 BY SIMILARITY.
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 788 788 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1034 1034 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1079 1079 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1121 1121 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1354 1354 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPPLIC 1072 1162 MISSING (IN MAJOR ISOFORM).
FT VARSPPLIC 1072 1253 MISSING (IN MINOR-1 ISOFORM).
FT CONFLICT 1007 1007 T -> M (IN REF. 2).
SQ SEQUENCE 1746 AA; 191399 MW; 56349B1CFE5E5C88 CRC64;

```

Query Match 12.3%; Score 139.5; DB 1; Length 1746;
 Best Local Similarity 26.2%; Pred. No. 0.00016;
 Matches 53; Conservative 17; Mismatches 69; Indels 63; Gaps 13;

```

OY 12 ANKTCPTNMMNNHIC--RCLAEDEPFMSSDAGDSDTG--FHDIQGNKELDEETCC 66
DB 248 SRETCPVPCSEHGRCVDRVCQEGF-----AGEDCNBPLCLHNGHGRVENE---C 299
OY 67 VCRAGLRAPASCG---PHKELDRNSC---QCYCKNKLFPSCGKANPEPENC----- 112
DB 300 VDEGFTGDEGELICPKCFPRRCINCTCYCEGFEDEG--RLACPHGGRGRCE 357
OY 113 --QCVC-----KRTCPRNQPLNPGKC--ACECTESPQCKLKKKFFHQTSCY 157
DB 358 EGCQVDEDFAGADCSERCP--SDCHNRGCLDGRCECDG-----FEGECGCEL 406
OY 158 RRP-----CTNRKACEPG 171
DB 407 RCPGCGSGHRCVNGQCVCEG 428

```

```

RESULT 8
NOTC_BRARE STANDARD: PRT: 2437 AA.
AC P46530:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN PRECURSOR.
GN NOTCH.
OS Brachydanio rerio (zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Rasbora; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA MEDLINE=94128602; PubMed=8297791;
RA Bierkamp C.; Campos-Ortega J.A.;

```

```

RT "A zebrafish homologue of the Drosophila neurogenic gene Notch and
RT its pattern of transcription during early embryogenesis.";
RL Mech. Dev. 43:87-100(1993).
CC -1- FUNCTION: IMPLICATED IN CELL FATE SPECIFICATIONS DURING
CC EMBRYO DEVELOPMENT. MAY BE INVOLVED IN THE FORMATION OF THE
CC NEURAL PLATE, NOTOCHORD AND BRAIN VESICLES.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN ALL CELLS IN PREGASTRULATION
CC STAGES. DURING GASTRULATION IS DIFFERENTIALLY EXPRESSED,
CC ACCUMULATING PREDOMINANTLY IN THE PRECHORDAL MESODERM AND
CC NOTOCHORD. AT THE END OF GASTRULATION, EXPRESSED ALONG THE
CC ANTERIOR-POSTERIOR AXIS INCLUDING THE DEVELOPING NEURAL PLATE
CC AND DIFFERENTIATING MESODERM. ALSO PRESENT IN THE DEVELOPING
CC BRAIN AND HEAD REGIONS.
CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -1- SIMILARITY: CONTAINS 6 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X69088; CAA48831.1; -.
DR HSSP: P00740; IIXA.
DR ZFIN: ZDB-GENE-990415-173; notch.
DR InterPro: IPR000152; -.
DR InterPro: IPR000561; -.
DR InterPro: IPR000800; -.
DR InterPro: IPR001336; -.
DR InterPro: IPR001438; -.
DR InterPro: IPR001881; -.
DR InterPro: IPR021110; -.
DR Pfam: PF00008; EGF; 36.
DR Pfam: PF00023; ank; 6.
DR Pfam: PF00066; notch; 3.
DR PRINTS: PR00009; EGFNGF.
DR PRINTS: PR00010; EGBLOOD.
DR PROSITE: PS50088; ANK_REPEAT; 4.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; 23.
DR PROSITE: PS00022; EGF_1; 34.
DR PROSITE: PS01186; EGF_2; 28.
DR PROSITE: PS01187; EGF_CA; 22.
KW Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;
KW Transmembrane; Signal; Glycoprotein.
FT SIGNAL 1 20
FT CHAIN 21 2437
FT DOMAIN 21 1724
FT TRANSMEM 1725 1747
FT DOMAIN 1748 2437
FT DOMAIN 21 57
FT DOMAIN 58 98
FT DOMAIN 101 138
FT DOMAIN 139 175
FT DOMAIN 177 215
FT DOMAIN 217 254
FT DOMAIN 256 292
FT DOMAIN 294 332
FT DOMAIN 334 370
FT DOMAIN 371 409
FT DOMAIN 411 449
FT DOMAIN 451 487
FT DOMAIN 489 524
FT DOMAIN 526 562
FT DOMAIN 564 599
FT DOMAIN 601 637
FT DOMAIN 639 674
FT DOMAIN 676 712

```

```

FT DOMAIN 714 749 EGF-LIKE 19, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 751 787 EGF-LIKE 20, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 789 825 EGF-LIKE 21, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 827 865 EGF-LIKE 22, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 867 903 EGF-LIKE 23, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 905 941 EGF-LIKE 24, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 943 979 EGF-LIKE 25, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 981 1017 EGF-LIKE 26, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1019 1055 EGF-LIKE 27, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1057 1093 EGF-LIKE 28, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1095 1141 EGF-LIKE 29, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1143 1179 EGF-LIKE 30, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1181 1217 EGF-LIKE 31, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1219 1263 EGF-LIKE 32, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1265 1303 EGF-LIKE 33, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1305 1344 EGF-LIKE 34, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1346 1382 EGF-LIKE 35, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1385 1423 EGF-LIKE 36, CALCIUM-BINDING (POTENTIAL).
FT REPEAT 1446 1486 LIN/NOTCH 1, CALCIUM-BINDING (POTENTIAL).
FT REPEAT 1487 1520 LIN/NOTCH 2, CALCIUM-BINDING (POTENTIAL).
FT REPEAT 1521 1561 LIN/NOTCH 3, CALCIUM-BINDING (POTENTIAL).
FT REPEAT 1867 1910 ANK 1, CALCIUM-BINDING (POTENTIAL).
FT REPEAT 1915 1944 ANK 2, CALCIUM-BINDING (POTENTIAL).
FT REPEAT 1948 1978 ANK 3, CALCIUM-BINDING (POTENTIAL).
FT REPEAT 1982 2011 ANK 4, CALCIUM-BINDING (POTENTIAL).
FT REPEAT 2015 2044 ANK 5, CALCIUM-BINDING (POTENTIAL).
FT REPEAT 2048 2077 ANK 6, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 2255 2276 POLY-GLN (OPR-REPEAT), CALCIUM-BINDING (POTENTIAL).
FT DISULFID 25 35 BY SIMILARITY.
FT DISULFID 29 45 BY SIMILARITY.
FT DISULFID 47 56 BY SIMILARITY.
FT DISULFID 62 73 BY SIMILARITY.
FT DISULFID 67 86 BY SIMILARITY.
FT DISULFID 88 97 BY SIMILARITY.
FT DISULFID 105 116 BY SIMILARITY.
FT DISULFID 110 126 BY SIMILARITY.
FT DISULFID 128 137 BY SIMILARITY.
FT DISULFID 143 154 BY SIMILARITY.
FT DISULFID 148 163 BY SIMILARITY.
FT DISULFID 165 174 BY SIMILARITY.
FT DISULFID 181 194 BY SIMILARITY.
FT DISULFID 188 203 BY SIMILARITY.
FT DISULFID 205 214 BY SIMILARITY.
FT DISULFID 221 232 BY SIMILARITY.
FT DISULFID 226 242 BY SIMILARITY.
FT DISULFID 244 253 BY SIMILARITY.
FT DISULFID 260 271 BY SIMILARITY.
FT DISULFID 265 280 BY SIMILARITY.
FT DISULFID 282 291 BY SIMILARITY.
FT DISULFID 298 311 BY SIMILARITY.
FT DISULFID 305 320 BY SIMILARITY.
FT DISULFID 322 331 BY SIMILARITY.
FT DISULFID 338 349 BY SIMILARITY.
FT DISULFID 343 358 BY SIMILARITY.
FT DISULFID 360 369 BY SIMILARITY.
FT DISULFID 375 386 BY SIMILARITY.
FT DISULFID 380 397 BY SIMILARITY.
FT DISULFID 399 408 BY SIMILARITY.
FT DISULFID 415 428 BY SIMILARITY.
FT DISULFID 422 437 BY SIMILARITY.
FT DISULFID 439 448 BY SIMILARITY.
FT DISULFID 455 466 BY SIMILARITY.
FT DISULFID 460 475 BY SIMILARITY.
FT DISULFID 477 486 BY SIMILARITY.
FT DISULFID 493 503 BY SIMILARITY.
FT DISULFID 498 512 BY SIMILARITY.
FT DISULFID 514 523 BY SIMILARITY.
FT DISULFID 530 541 BY SIMILARITY.
FT DISULFID 535 550 BY SIMILARITY.
FT DISULFID 552 561 BY SIMILARITY.
FT DISULFID 568 578 BY SIMILARITY.
FT DISULFID 573 587 BY SIMILARITY.
FT DISULFID 589 598 BY SIMILARITY.

```

```

FT DISULFID 605 616 BY SIMILARITY.
FT DISULFID 610 625 BY SIMILARITY.
FT DISULFID 627 636 BY SIMILARITY.
FT DISULFID 643 653 BY SIMILARITY.
FT DISULFID 648 662 BY SIMILARITY.
FT DISULFID 664 673 BY SIMILARITY.
FT DISULFID 680 691 BY SIMILARITY.
FT DISULFID 691 700 BY SIMILARITY.
FT DISULFID 702 711 BY SIMILARITY.
FT DISULFID 718 728 BY SIMILARITY.
FT DISULFID 723 737 BY SIMILARITY.
FT DISULFID 739 748 BY SIMILARITY.
FT DISULFID 755 766 BY SIMILARITY.
FT DISULFID 760 775 BY SIMILARITY.
FT DISULFID 777 786 BY SIMILARITY.
FT DISULFID 793 804 BY SIMILARITY.
FT DISULFID 798 813 BY SIMILARITY.
FT DISULFID 815 824 BY SIMILARITY.
FT DISULFID 831 842 BY SIMILARITY.
FT DISULFID 836 853 BY SIMILARITY.
FT DISULFID 855 864 BY SIMILARITY.
FT DISULFID 871 882 BY SIMILARITY.
FT DISULFID 876 891 BY SIMILARITY.
FT DISULFID 893 902 BY SIMILARITY.
FT DISULFID 909 920 BY SIMILARITY.
FT DISULFID 914 929 BY SIMILARITY.
FT DISULFID 931 940 BY SIMILARITY.
FT DISULFID 947 958 BY SIMILARITY.
FT DISULFID 952 967 BY SIMILARITY.
FT DISULFID 969 978 BY SIMILARITY.
FT DISULFID 1023 1034 BY SIMILARITY.
FT DISULFID 1028 1043 BY SIMILARITY.
FT DISULFID 1045 1054 BY SIMILARITY.
FT DISULFID 1061 1072 BY SIMILARITY.
FT DISULFID 1066 1081 BY SIMILARITY.
FT DISULFID 1083 1092 BY SIMILARITY.
FT DISULFID 1099 1120 BY SIMILARITY.

```

Query Match 12.2%; Score 137.5; DB 1; Length 2437;

Best Local Similarity 22.9%; Pred. No. 0.00031; Matches 57; Conservative 13; Mismatches 66; Indels 113; Gaps 15;

```

QY 44 DSTDGFHDCGP-----NKEIDE-----ETC-----OCVCRAGLRPASC-GP 79
DB 655 DKINGECCECEGYSSMKNINIDCALNPGCHNGCIDGVNSFTCLCPDGRFADATCLSO 714
QY 80 HKELDRNSC-----OCVCKNKLFPSCCGAN-REPDENFC-----Q 113
DB 715 HNECSNPICIHGSCLDQINSYRCVCEAGMGRNCDININECLSNPCVNGTCKDMTSGYL 774
QY 114 CYCK-----RCPRNQPINPGKC-----ACECTE-----SPQ 140
DB 775 CTRAGFSGPNCQMINNECASNPCCLNQGSCIDVAGFKNCMLPYTGECVENLAPCSPR 834
QY 141 KCLLKG-----KRFHQFC-----SCYRRPCTNROKACE---PGFSYSE 176
DB 835 PCKNGGVCESEDPQSFSCNCPAGNGGTCEVDINECVANPCPN-GGVENLNRGGFO--- 890
QY 177 EYCRVPSY 185
DB 891 --CRCNPGF 897

```

```

RESULT 9
NOTC_XENLA
ID NOTC_XENLA STANDARD; PRT; 2524 AA.
AC P21783;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG PRECURSOR (XOTCH PROTEIN).
GN XOTCH.
OS Xenopus laevis (African clawed frog).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=90385285; PubMed=2402639;
RX Coffman C., Harris W., Kintner C.;
RT "Notch, the Xenopus homolog of Drosophila notch.";
RL Science 249:1438-1441(1990).
RN [2]
RP REVISIONS TO 1759-1782.
RA Kintner C.;
RL Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED ALMOST UNIFORMLY IN EARLY EMBRYOS.
CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
CC -1- SIMILARITY: CONTRAINS 36 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTRAINS 3 LIN/NOTCH REPEATS.
CC -1- SIMILARITY: CONTRAINS 6 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL, M33674; AAB02039.1; -
DR PIR, A35844; A35844.
DR HSSP, P00740; 11XA.
DR InterPro: IPR000152; -
DR InterPro: IPR000561; -
DR InterPro: IPR000800; -
DR InterPro: IPR001438; -
DR InterPro: IPR001881; -
DR InterPro: IPR002110; -
DR Pfam: PF00008; EGF; 36.
DR Pfam: PF00023; notch; 3.
DR Pfam: PF00066; notch; 3.
DR PRINTS: PR00010; EGBLOOD.
DR PROSITE: PS5008; ANK_REPEAT; 4.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; 23.
DR PROSITE: PS00022; EGF_1; 34.
DR PROSITE: PS01186; EGF_2; 29.
DR PROSITE: PS01187; EGF_Ca; 21.
KM Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;
KW Transmembrane; Signal; Glycoprotein.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 2524 NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG.
FT DOMAIN 20 1728 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1729 1750 POTENTIAL.
FT DOMAIN 1751 2524 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 20 57 EGF-LIKE 1.
FT DOMAIN 58 99 EGF-LIKE 2.
FT DOMAIN 102 140 EGF-LIKE 3.
FT DOMAIN 141 177 EGF-LIKE 4.
FT DOMAIN 179 215 EGF-LIKE 5.
FT DOMAIN 217 254 EGF-LIKE 6.
FT DOMAIN 256 292 EGF-LIKE 7.
FT DOMAIN 294 332 EGF-LIKE 8.
FT DOMAIN 334 370 EGF-LIKE 9.
FT DOMAIN 371 409 EGF-LIKE 10.
FT DOMAIN 411 449 EGF-LIKE 11.
FT DOMAIN 451 487 EGF-LIKE 12.
FT DOMAIN 489 525 EGF-LIKE 13.
FT DOMAIN 527 563 EGF-LIKE 14.
FT DOMAIN 565 600 EGF-LIKE 15.
FT DOMAIN 602 638 EGF-LIKE 16.
FT DOMAIN 640 675 EGF-LIKE 17.
FT DOMAIN 677 713 EGF-LIKE 18.
FT DOMAIN 713 750 EGF-LIKE 19.
FT DOMAIN 752 788 EGF-LIKE 20.
FT DOMAIN 790 826 EGF-LIKE 21.
FT DOMAIN 828 866 EGF-LIKE 22.
FT DOMAIN 868 904 EGF-LIKE 23.
FT DOMAIN 906 942 EGF-LIKE 24.
FT DOMAIN 944 980 EGF-LIKE 25.
FT DOMAIN 982 1018 EGF-LIKE 26.
FT DOMAIN 1020 1056 EGF-LIKE 27.
FT DOMAIN 1058 1094 EGF-LIKE 28.
FT DOMAIN 1096 1142 EGF-LIKE 29.
FT DOMAIN 1144 1180 EGF-LIKE 30.
FT DOMAIN 1182 1218 EGF-LIKE 31.
FT DOMAIN 1220 1264 EGF-LIKE 32.
FT DOMAIN 1266 1304 EGF-LIKE 33.
FT DOMAIN 1306 1346 EGF-LIKE 34.
FT DOMAIN 1347 1383 EGF-LIKE 35.
FT DOMAIN 1386 1424 EGF-LIKE 36.
FT DOMAIN 1424 1478 LIN/NOTCH 1.
FT DOMAIN 1479 1520 LIN/NOTCH 2.
FT REPEAT 1521 1560 LIN/NOTCH 3.
FT REPEAT 1560 1599 ANK 1.
FT REPEAT 1599 1637 ANK 2.
FT REPEAT 1637 1675 ANK 3.
FT REPEAT 1675 1713 ANK 4.
FT REPEAT 1713 1751 ANK 5.
FT REPEAT 1751 1789 ANK 6.
FT DISULFID 29 45 BY SIMILARITY.
FT DISULFID 47 56 BY SIMILARITY.
FT DISULFID 62 74 BY SIMILARITY.
FT DISULFID 68 87 BY SIMILARITY.
FT DISULFID 89 98 BY SIMILARITY.
FT DISULFID 106 117 BY SIMILARITY.
FT DISULFID 111 128 BY SIMILARITY.
FT DISULFID 130 139 BY SIMILARITY.
FT DISULFID 145 156 BY SIMILARITY.
FT DISULFID 150 165 BY SIMILARITY.
FT DISULFID 167 176 BY SIMILARITY.
FT DISULFID 183 194 BY SIMILARITY.
FT DISULFID 188 203 BY SIMILARITY.
FT DISULFID 205 214 BY SIMILARITY.
FT DISULFID 221 232 BY SIMILARITY.
FT DISULFID 226 242 BY SIMILARITY.
FT DISULFID 244 253 BY SIMILARITY.
FT DISULFID 260 271 BY SIMILARITY.
FT DISULFID 265 280 BY SIMILARITY.
FT DISULFID 282 291 BY SIMILARITY.
FT DISULFID 298 311 BY SIMILARITY.
FT DISULFID 305 320 BY SIMILARITY.
FT DISULFID 322 331 BY SIMILARITY.
FT DISULFID 338 349 BY SIMILARITY.
FT DISULFID 343 358 BY SIMILARITY.
FT DISULFID 360 369 BY SIMILARITY.
FT DISULFID 375 386 BY SIMILARITY.
FT DISULFID 380 397 BY SIMILARITY.
FT DISULFID 399 408 BY SIMILARITY.
FT DISULFID 415 428 BY SIMILARITY.
FT DISULFID 422 437 BY SIMILARITY.
FT DISULFID 439 448 BY SIMILARITY.
FT DISULFID 455 466 BY SIMILARITY.
FT DISULFID 477 486 BY SIMILARITY.
FT DISULFID 493 504 BY SIMILARITY.
FT DISULFID 498 513 BY SIMILARITY.
FT DISULFID 515 524 BY SIMILARITY.
FT DISULFID 531 542 BY SIMILARITY.
FT DISULFID 542 551 BY SIMILARITY.
FT DISULFID 553 562 BY SIMILARITY.
FT DISULFID 569 579 BY SIMILARITY.
FT DISULFID 574 588 BY SIMILARITY.
FT DISULFID 590 599 BY SIMILARITY.
FT DISULFID 606 617 BY SIMILARITY.

```

FT DISULFD 611 626 BY SIMILARITY.
FT DISULFD 628 637 BY SIMILARITY.
FT DISULFD 644 654 BY SIMILARITY.
FT DISULFD 649 663 BY SIMILARITY.
FT DISULFD 665 674 BY SIMILARITY.
FT DISULFD 681 692 BY SIMILARITY.
FT DISULFD 686 701 BY SIMILARITY.
FT DISULFD 703 712 BY SIMILARITY.
FT DISULFD 719 729 BY SIMILARITY.
FT DISULFD 724 738 BY SIMILARITY.
FT DISULFD 740 749 BY SIMILARITY.
FT DISULFD 756 767 BY SIMILARITY.
FT DISULFD 761 776 BY SIMILARITY.
FT DISULFD 778 787 BY SIMILARITY.
FT DISULFD 794 805 BY SIMILARITY.
FT DISULFD 799 814 BY SIMILARITY.
FT DISULFD 816 825 BY SIMILARITY.
FT DISULFD 832 843 BY SIMILARITY.
FT DISULFD 837 854 BY SIMILARITY.
FT DISULFD 856 865 BY SIMILARITY.
FT DISULFD 872 883 BY SIMILARITY.
FT DISULFD 877 892 BY SIMILARITY.
FT DISULFD 894 903 BY SIMILARITY.
FT DISULFD 910 921 BY SIMILARITY.
FT DISULFD 915 930 BY SIMILARITY.
FT DISULFD 932 941 BY SIMILARITY.
FT DISULFD 966 997 BY SIMILARITY.
FT DISULFD 991 1006 BY SIMILARITY.
FT DISULFD 1008 1017 BY SIMILARITY.
FT DISULFD 1024 1035 BY SIMILARITY.
FT DISULFD 1029 1044 BY SIMILARITY.
FT DISULFD 1046 1055 BY SIMILARITY.
FT DISULFD 1062 1073 BY SIMILARITY.
FT DISULFD 1067 1082 BY SIMILARITY.
FT DISULFD 1084 1093 BY SIMILARITY.
FT DISULFD 1100 1121 BY SIMILARITY.
FT DISULFD 1115 1130 BY SIMILARITY.
FT DISULFD 1132 1141 BY SIMILARITY.
FT DISULFD 1148 1159 BY SIMILARITY.
FT DISULFD 1153 1168 BY SIMILARITY.
FT DISULFD 1170 1179 BY SIMILARITY.
FT DISULFD 1186 1197 BY SIMILARITY.
FT DISULFD 1191 1206 BY SIMILARITY.
FT DISULFD 1208 1217 BY SIMILARITY.
FT DISULFD 1224 1243 BY SIMILARITY.
FT DISULFD 1237 1252 BY SIMILARITY.

```

Query Match 12.2%, Score 137.5; DB 1; Length 2524;
 Best Local Similarity 21.7%; Pred. No. 0.00032;
 Matches 55; Conservative 21; Mismatches 77; Indels 101; Gaps 12;

```

QY 7 POCQANTCPNNYNNMNN-----HICRCL-----AQEDFMSSDAGDSTGCFHDI 52
DB 765 POCQNTNINECSSNPLNCTCIDVAGKCCNMCMLPYTAICICAVIAPCAGSPCKNG---- 840
QY 53 CGPNKEDE-ETQCQVCRAGLPASCGRPHKEIDRNSC-----QCVCCK 93
DB 841 -GRCKESDPEFFSECECPGNGGTC-----EIDMNECVNRPCRNATGNTGSKCKCK 895
QY 94 NKLPSQCGANREPDENTCQCVCARTCPRNOLNPGKCA-----CECTESPOKCLKG 146
DB 896 ----PGYGRNCNEMDIDCO-----PNCPHNGSCSDGINMFCMCAPAG-----FRG 938
QY 147 KKEHHQTCSCYRRPCTNRK-----ACEPES----- 173
DB 939 PCCEEDINECSNPKCNKANCNTCVNSTYTCQPGFSGIHCSNTPDCTESSCFNGGTCT 998
QY 174 --YSEEVCRCPVSY 185
DB 999 DGINFCTCCPGPF 1012

```

```

FBN2_MOUSE
ID FBN2_MOUSE STANDARD; PRT: 2907 AA.
AC Q61555; Q63957;
DT 15-DEC-1998 (rel. 37, Created)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 15-DEC-1998 (rel. 37, Last annotation update)
DE FIBRILLIN 2 PRECURSOR.
GN FBN2 OR FBN-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95263670; PubMed=7744963;
RA Zhang H., Hu W., Ramirez F.;
RT "Developmental expression of fibrillin genes suggests heterogeneity
of extracellular microfibrils."
RL J. Cell Biol. 129:1165-1176(1995).
RN [2]
RP SEQUENCE OF 210-317 FROM N.A.
RX MEDLINE=94140368; PubMed=8307578;
RA Li X., Pereira L., Zhang H., Sanguinetti C., Ramirez F., Bonadio J.,
RA Francke U.;
RT "Fibrillin genes map to regions of conserved mouse/human synteny on
mouse chromosomes 2 and 18."
RL Genomics 18:667-672(1993).
CC -I- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC -I- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
EGF-LIKE DOMAINS.
CC -I- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sdb.ch/announce/
or send an email to license@sdb-sdb.ch).
CC -----
DR EMBL: L39790; AAA74908.1;
DR EMBL: S69359; AAC60685.1;
DR MGD: MGI:95490; Fbn2.
DR InterPro: IPR000152;
DR InterPro: IPR000561;
DR InterPro: IPR000822;
DR InterPro: IPR001438;
DR InterPro: IPR001881;
DR InterPro: IPR00212;
DR Pfam: PF00008; EGF; 46.
DR Pfam: PF00683; TB; 9.
DR PRINTS: PR00010; EGFBLD.
DR PROSITE: PS00010; ASX_HYDROXYL; 43.
DR PROSITE: PS00022; ASX; 1; 2.
DR PROSITE: PS01186; EGF 2; 36.
DR PROSITE: PS01187; EGF_CA; 43.
KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW Repeat; Signal; Multigene family.
FT SIGNAL 1 28
FT CHAIN 29 2907
FT DOMAIN 111 142 FIBRILLIN 2.
FT DOMAIN 145 176 EGF-LIKE 1, NON-CALCIUM BINDING.
FT DOMAIN 176 176 EGF-LIKE 2, NON-CALCIUM BINDING.
FT DOMAIN 176 208 EGF-LIKE 3, NON-CALCIUM BINDING.
FT DOMAIN 276 317 EGF-LIKE 4, CALCIUM-BINDING.
FT DOMAIN 318 359 EGF-LIKE 5, CALCIUM-BINDING.
FT REPEAT 360 426 TGFBP 1.
FT DOMAIN 487 527 EGF-LIKE 6, NON-CALCIUM BINDING.
FT DOMAIN 528 567 EGF-LIKE 7, CALCIUM-BINDING.
FT DOMAIN 568 609 EGF-LIKE 8, CALCIUM-BINDING.
FT DOMAIN 610 650 EGF-LIKE 9, CALCIUM-BINDING.
FT DOMAIN 651 691 EGF-LIKE 10, CALCIUM-BINDING.

```

```
FT REPEAT 692 760 TGFBP 2.
FT DOMAIN 761 802 EGF-LIKE 11, CALCIUM-BINDING.
FT DOMAIN 803 844 EGF-LIKE 12, CALCIUM-BINDING.
FT DOMAIN 845 883 EGF-LIKE 13, CALCIUM-BINDING.
FT DOMAIN 884 905 EGF-LIKE 14, CALCIUM-BINDING.
FT REPEAT 906 1065 TGFBP 3.
FT DOMAIN 1066 1107 EGF-LIKE 15, CALCIUM-BINDING.
FT DOMAIN 1108 1150 EGF-LIKE 16, CALCIUM-BINDING.
FT DOMAIN 1151 1192 EGF-LIKE 17, CALCIUM-BINDING.
FT DOMAIN 1193 1234 EGF-LIKE 18, CALCIUM-BINDING.
FT DOMAIN 1235 1275 EGF-LIKE 19, CALCIUM-BINDING.
FT DOMAIN 1276 1317 EGF-LIKE 20, CALCIUM-BINDING.
FT DOMAIN 1318 1359 EGF-LIKE 21, CALCIUM-BINDING.
FT DOMAIN 1360 1400 EGF-LIKE 22, CALCIUM-BINDING.
FT DOMAIN 1401 1441 EGF-LIKE 23, CALCIUM-BINDING.
FT DOMAIN 1442 1483 EGF-LIKE 24, CALCIUM-BINDING.
FT DOMAIN 1484 1524 EGF-LIKE 25, CALCIUM-BINDING.
FT DOMAIN 1525 1565 EGF-LIKE 26, CALCIUM-BINDING.
FT REPEAT 1566 1642 TGFBP 4.
FT DOMAIN 1643 1684 EGF-LIKE 27, CALCIUM-BINDING.
FT REPEAT 1685 1726 EGF-LIKE 28, CALCIUM-BINDING.
FT REPEAT 1727 1800 TGFBP 5.
FT DOMAIN 1801 1842 EGF-LIKE 29, CALCIUM-BINDING.
FT DOMAIN 1843 1884 EGF-LIKE 30, CALCIUM-BINDING.
FT DOMAIN 1885 1926 EGF-LIKE 31, CALCIUM-BINDING.
FT DOMAIN 1927 1965 EGF-LIKE 32, CALCIUM-BINDING.
FT DOMAIN 1966 2008 EGF-LIKE 33, CALCIUM-BINDING.
FT DOMAIN 2009 2048 EGF-LIKE 34, CALCIUM-BINDING.
FT DOMAIN 2049 2090 EGF-LIKE 35, CALCIUM-BINDING.
FT REPEAT 2091 2163 TGFBP 6.
FT DOMAIN 2164 2205 EGF-LIKE 36, CALCIUM-BINDING.
FT DOMAIN 2206 2245 EGF-LIKE 37, CALCIUM-BINDING.
FT DOMAIN 2246 2286 EGF-LIKE 38, CALCIUM-BINDING.
FT DOMAIN 2287 2330 EGF-LIKE 39, CALCIUM-BINDING.
FT DOMAIN 2331 2372 EGF-LIKE 40, CALCIUM-BINDING.
FT REPEAT 2373 2441 TGFBP 7.
FT DOMAIN 2442 2483 EGF-LIKE 41, CALCIUM-BINDING.
FT DOMAIN 2484 2524 EGF-LIKE 42, CALCIUM-BINDING.
FT DOMAIN 2525 2563 EGF-LIKE 43, CALCIUM-BINDING.
FT DOMAIN 2564 2606 EGF-LIKE 44, CALCIUM-BINDING.
FT DOMAIN 2607 2646 EGF-LIKE 45, CALCIUM-BINDING.
FT DOMAIN 2647 2687 EGF-LIKE 46, CALCIUM-BINDING.
FT DOMAIN 2688 2727 EGF-LIKE 47, CALCIUM-BINDING.
FT DISULFID 115 124 BY SIMILARITY.
FT DISULFID 119 130 BY SIMILARITY.
FT DISULFID 132 141 BY SIMILARITY.
FT DISULFID 149 159 BY SIMILARITY.
FT DISULFID 153 164 BY SIMILARITY.
FT DISULFID 166 175 BY SIMILARITY.
FT DISULFID 180 190 BY SIMILARITY.
FT DISULFID 184 196 BY SIMILARITY.
FT DISULFID 198 207 BY SIMILARITY.
FT DISULFID 280 292 BY SIMILARITY.
FT DISULFID 287 301 BY SIMILARITY.
FT DISULFID 303 316 BY SIMILARITY.
FT DISULFID 322 334 BY SIMILARITY.
FT DISULFID 329 343 BY SIMILARITY.
FT DISULFID 345 358 BY SIMILARITY.
FT DISULFID 491 503 BY SIMILARITY.
FT DISULFID 498 512 BY SIMILARITY.
FT DISULFID 514 526 BY SIMILARITY.
FT DISULFID 532 542 BY SIMILARITY.
FT DISULFID 537 551 BY SIMILARITY.
FT DISULFID 553 566 BY SIMILARITY.
FT DISULFID 572 584 BY SIMILARITY.
FT DISULFID 579 593 BY SIMILARITY.
FT DISULFID 595 608 BY SIMILARITY.
FT DISULFID 614 625 BY SIMILARITY.
FT DISULFID 620 634 BY SIMILARITY.
FT DISULFID 636 649 BY SIMILARITY.
FT DISULFID 655 666 BY SIMILARITY.
FT DISULFID 661 675 BY SIMILARITY.
FT DISULFID 677 690 BY SIMILARITY.
```

```
FT DISULFID 765 777 BY SIMILARITY.
FT DISULFID 772 786 BY SIMILARITY.
FT DISULFID 788 801 BY SIMILARITY.
FT DISULFID 807 819 BY SIMILARITY.
FT DISULFID 814 828 BY SIMILARITY.
FT DISULFID 830 843 BY SIMILARITY.
FT DISULFID 849 859 BY SIMILARITY.
FT DISULFID 854 868 BY SIMILARITY.
FT DISULFID 870 883 BY SIMILARITY.
FT DISULFID 952 964 BY SIMILARITY.
FT DISULFID 959 973 BY SIMILARITY.
FT DISULFID 975 988 BY SIMILARITY.
FT DISULFID 1070 1082 BY SIMILARITY.
FT DISULFID 1077 1091 BY SIMILARITY.
FT DISULFID 1093 1106 BY SIMILARITY.
FT DISULFID 1112 1124 BY SIMILARITY.
FT DISULFID 1119 1133 BY SIMILARITY.
FT DISULFID 1135 1149 BY SIMILARITY.
FT DISULFID 1155 1167 BY SIMILARITY.
FT DISULFID 1162 1176 BY SIMILARITY.
FT DISULFID 1178 1191 BY SIMILARITY.
FT DISULFID 1197 1209 BY SIMILARITY.
FT DISULFID 1204 1218 BY SIMILARITY.
FT DISULFID 1220 1233 BY SIMILARITY.
FT DISULFID 1239 1250 BY SIMILARITY.
FT DISULFID 1246 1259 BY SIMILARITY.
FT DISULFID 1261 1274 BY SIMILARITY.
FT DISULFID 1280 1292 BY SIMILARITY.
FT DISULFID 1287 1301 BY SIMILARITY.
FT DISULFID 1303 1316 BY SIMILARITY.
FT DISULFID 1332 1343 BY SIMILARITY.
FT DISULFID 1329 1343 BY SIMILARITY.
FT DISULFID 1345 1358 BY SIMILARITY.
FT DISULFID 1354 1377 BY SIMILARITY.
FT DISULFID 1371 1386 BY SIMILARITY.
FT DISULFID 1388 1399 BY SIMILARITY.
FT DISULFID 1405 1418 BY SIMILARITY.
FT DISULFID 1412 1427 BY SIMILARITY.
FT DISULFID 1429 1440 BY SIMILARITY.
FT DISULFID 1446 1458 BY SIMILARITY.
FT DISULFID 1453 1467 BY SIMILARITY.
FT DISULFID 1469 1482 BY SIMILARITY.
FT DISULFID 1488 1499 BY SIMILARITY.
FT DISULFID 1494 1508 BY SIMILARITY.
FT DISULFID 1510 1523 BY SIMILARITY.
FT DISULFID 1529 1540 BY SIMILARITY.
FT DISULFID 1535 1549 BY SIMILARITY.
FT DISULFID 1551 1564 BY SIMILARITY.
FT DISULFID 1647 1659 BY SIMILARITY.
FT DISULFID 1654 1668 BY SIMILARITY.
FT DISULFID 1670 1683 BY SIMILARITY.
FT DISULFID 1689 1701 BY SIMILARITY.
FT DISULFID 1696 1710 BY SIMILARITY.
FT DISULFID 1712 1725 BY SIMILARITY.
FT DISULFID 1805 1817 BY SIMILARITY.
```

Query Match 12.0%: Score 136; DB 1; Length 2907;

Best Local Similarity 23.9%: Pred. No. 0.00048; Matches 60; Conservative 24; Mismatches 83; Indels 84; Gaps 15;

```
QY 6 LPQCOANKTCPTNYMNN---HICRC---LAQEDFMFSSDAGDDST----- 46
DB 2485 LDECSQSPKPC--NFIKNTKGSYQCSPPRGVILQEDGKTKDDEGCTKHNNQFLCVN 2542
QY 47 -----DGF---HICGPNKELDEE-----TCQ-----CVCRAGL-----RP 74
DB 2543 TLGGFTCKCPPTGHTNACIDNNNGSGSPSLCGAKGICQNTPGSFFSCQGFSLDASG 2602
QY 75 ASCGPNKELDRN--SCQCVCKNKLPSSQCGA-----NRPEDNTCCGVCVKRTCPRN 123
DB 2603 LNCELYVDECDGNHRCQGNQNLILGGYRCGCFRNGDYONHONQOCVDENSCS----- 2652
QY 124 QPLNPGKC---ACECTESPQCLD--KGRKFHHQTCSCYR--RPTNROKACEPFGSYSE-- 176
```



```

Db 2653 ---NPGACGASACVNTGSGKACPGSFSDQFSSACHDVNCCSSKNCPCSYGCSNTEGG 2709
OY 177 EVCRCVPSYWK 187
Db 2710 YLGGCPGTYR 2720

RESULT 11
FBN2_HUMAN
ID FBN2_HUMAN STANDARD; PRT; 2911 AA.
AC P35556;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE FIBRILLIN 2 PRECURSOR.
GN FBN2
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94165150; PubMed=8120105;
RA Zhang H., Apeltroth S.D., Hu W., Davis E.C., Sanguineti C.,
RA Bonadio J., Mechem R.P., Ramirez F.;
RT "Structure and expression of fibrillin-2, a novel microfibrillar
RT component preferentially located in elastic matrices.";
RL J. Cell Biol. 124:855-863(1994).
RN [2]
RP SEQUENCE OF 752-1505 FROM N.A.
RX MEDLINE=91304567; PubMed=1852206;
RA Lee B., Godfrey M., Vitale E., Horl H., Mattei M.-G., Sarfarazi M.,
RA Tsipouras P., Ramirez F., Hollister D.;
RT "Linkage of Marfan syndrome and a phenotypically related disorder to
RT two different fibrillin genes.";
RL Nature 352:330-334(1991).
RN [3]
RP VARIANTS CCA TYR-1252 AND SER-1433, AND VARIANT ILE-964.
RX MEDLINE=96083599; PubMed=7493032;
RA Putnam E.A., Zhang H., Ramirez F., Milewicz D.M.;
RT "Fibrillin-2 (FBN2) mutations result in the Marfan-like disorder,
RT congenital contractural arachnodactyly.";
RL Nat. Genet. 11:456-458(1995).
CC -1- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC THAT BINDS CALCIUM. FIBRILLIN-2-CONTAINING MICROFIBRILS REGULATE
CC THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.
CC -1- DISEASE: DEFECTS IN FBN2 ARE THE CAUSE OF CONGENITAL CONTRACTURAL
CC ARACHNOACTYLY (CCA) (ALSO KNOWN AS BEALS SYNDROME). CCA IS
CC PHENOTYPICALLY SIMILAR TO MARFAN SYNDROME, BUT DOES NOT EFFECT THE
CC AORTA AND THE EYES.
CC -1- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
CC EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U03372; AAA18950.1;
CC EMBL: X62009; -, NOT_ANNOTATED_CDS.
CC PIR: S17063; S17063.
CC PIR: S31101; S31101.
CC HSSP: P35555; IEMO.
CC MIM: 121050;
CC InterPro: IPR000152;
CC InterPro: IPR000561;
CC InterPro: IPR001438;
CC InterPro: IPR001881;

```

```

DR InterPro: IPR002212;
DR Pfam: PF00008; EGF; 46.
DR Pfam: PF00683; TB; 9.
DR PRINTS: PR00010; EGFBLD.
DR PROSITE: PS00010; ASX_HYDROXYL; 43.
DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS01186; EGF_2; 37.
DR PROSITE: PS01187; EGF_CA; 43.
DR Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW Repeat; Signal; Multigene family; Disease mutation; Polymorphism.
FT SIGNAL 1
FT CHAIN 29
FT FIBRILLIN 2.
FT DOMAIN 111 142 EGF-LIKE 1, NON-CALCIUM BINDING.
FT DOMAIN 145 176 EGF-LIKE 2, NON-CALCIUM BINDING.
FT DOMAIN 176 207 EGF-LIKE 3, NON-CALCIUM BINDING.
FT DOMAIN 275 316 EGF-LIKE 4, CALCIUM-BINDING.
FT DOMAIN 317 358 EGF-LIKE 5, CALCIUM-BINDING.
FT REPEAT 359 425 TGFBP 1.
FT DOMAIN 493 533 EGF-LIKE 6, NON-CALCIUM BINDING.
FT DOMAIN 534 573 EGF-LIKE 7, CALCIUM-BINDING.
FT DOMAIN 574 615 EGF-LIKE 8, CALCIUM-BINDING.
FT DOMAIN 616 656 EGF-LIKE*9, CALCIUM-BINDING.
FT DOMAIN 657 697 EGF-LIKE 10, CALCIUM-BINDING.
FT REPEAT 698 766 TGFBP 2.
FT REPEAT 767 808 EGF-LIKE 11, CALCIUM-BINDING.
FT DOMAIN 809 850 EGF-LIKE 12, CALCIUM-BINDING.
FT DOMAIN 851 890 EGF-LIKE 13, CALCIUM-BINDING.
FT DOMAIN 954 995 EGF-LIKE 14, CALCIUM-BINDING.
FT REPEAT 1071 1071 TGFBP 3.
FT DOMAIN 1072 1113 EGF-LIKE 15, CALCIUM-BINDING.
FT DOMAIN 1114 1156 EGF-LIKE 16, CALCIUM-BINDING.
FT DOMAIN 1157 1198 EGF-LIKE 17, CALCIUM-BINDING.
FT DOMAIN 1199 1240 EGF-LIKE 18, CALCIUM-BINDING.
FT DOMAIN 1241 1281 EGF-LIKE 19, CALCIUM-BINDING.
FT DOMAIN 1282 1323 EGF-LIKE 20, CALCIUM-BINDING.
FT DOMAIN 1324 1365 EGF-LIKE 21, CALCIUM-BINDING.
FT DOMAIN 1366 1406 EGF-LIKE 22, CALCIUM-BINDING.
FT DOMAIN 1407 1447 EGF-LIKE 23, CALCIUM-BINDING.
FT DOMAIN 1448 1489 EGF-LIKE 24, CALCIUM-BINDING.
FT DOMAIN 1490 1530 EGF-LIKE 25, CALCIUM-BINDING.
FT DOMAIN 1531 1571 EGF-LIKE 26, CALCIUM-BINDING.
FT REPEAT 1572 1648 TGFBP 4.
FT DOMAIN 1649 1690 EGF-LIKE 27, CALCIUM-BINDING.
FT DOMAIN 1691 1732 EGF-LIKE 28, CALCIUM-BINDING.
FT REPEAT 1733 1806 TGFBP 5.
FT DOMAIN 1807 1848 EGF-LIKE 29, CALCIUM-BINDING.
FT DOMAIN 1849 1890 EGF-LIKE 30, CALCIUM-BINDING.
FT DOMAIN 1891 1932 EGF-LIKE 31, CALCIUM-BINDING.
FT DOMAIN 1933 1971 EGF-LIKE 32, CALCIUM-BINDING.
FT DOMAIN 1972 2014 EGF-LIKE 33, CALCIUM-BINDING.
FT DOMAIN 2015 2054 EGF-LIKE 34, CALCIUM-BINDING.
FT DOMAIN 2055 2096 EGF-LIKE 35, CALCIUM-BINDING.
FT REPEAT 2097 2169 TGFBP 6.
FT DOMAIN 2170 2211 EGF-LIKE 36, CALCIUM-BINDING.
FT DOMAIN 2212 2251 EGF-LIKE 37, CALCIUM-BINDING.
FT DOMAIN 2252 2292 EGF-LIKE 38, CALCIUM-BINDING.
FT DOMAIN 2293 2336 EGF-LIKE 39, CALCIUM-BINDING.
FT DOMAIN 2337 2378 EGF-LIKE 40, CALCIUM-BINDING.
FT REPEAT 2379 2447 TGFBP 7.
FT DOMAIN 2448 2489 EGF-LIKE 41, CALCIUM-BINDING.
FT DOMAIN 2490 2530 EGF-LIKE 42, CALCIUM-BINDING.
FT DOMAIN 2531 2569 EGF-LIKE 43, CALCIUM-BINDING.
FT DOMAIN 2570 2612 EGF-LIKE 44, CALCIUM-BINDING.
FT DOMAIN 2613 2652 EGF-LIKE 45, CALCIUM-BINDING.
FT DOMAIN 2653 2693 EGF-LIKE 46, CALCIUM-BINDING.
FT DOMAIN 2694 2733 EGF-LIKE 47, CALCIUM-BINDING.
FT DISULFD 115 124 BY SIMILARITY.
FT DISULFD 119 130 BY SIMILARITY.
FT DISULFD 132 141 BY SIMILARITY.
FT DISULFD 149 159 BY SIMILARITY.
FT DISULFD 153 164 BY SIMILARITY.
FT DISULFD 166 175 BY SIMILARITY.
FT DISULFD 180 190 BY SIMILARITY.

```

```
FT DISULFID 184 195 BY SIMILARITY.
FT DISULFID 197 206 BY SIMILARITY.
FT DISULFID 279 291 BY SIMILARITY.
FT DISULFID 286 300 BY SIMILARITY.
FT DISULFID 302 315 BY SIMILARITY.
FT DISULFID 321 333 BY SIMILARITY.
FT DISULFID 328 342 BY SIMILARITY.
FT DISULFID 344 357 BY SIMILARITY.
FT DISULFID 497 509 BY SIMILARITY.
FT DISULFID 504 518 BY SIMILARITY.
FT DISULFID 520 532 BY SIMILARITY.
FT DISULFID 538 548 BY SIMILARITY.
FT DISULFID 543 557 BY SIMILARITY.
FT DISULFID 559 572 BY SIMILARITY.
FT DISULFID 578 590 BY SIMILARITY.
FT DISULFID 585 599 BY SIMILARITY.
FT DISULFID 601 614 BY SIMILARITY.
FT DISULFID 620 631 BY SIMILARITY.
FT DISULFID 626 640 BY SIMILARITY.
FT DISULFID 642 655 BY SIMILARITY.
FT DISULFID 661 672 BY SIMILARITY.
FT DISULFID 667 681 BY SIMILARITY.
FT DISULFID 683 696 BY SIMILARITY.
FT DISULFID 771 783 BY SIMILARITY.
FT DISULFID 778 792 BY SIMILARITY.
FT DISULFID 794 807 BY SIMILARITY.
FT DISULFID 825 834 BY SIMILARITY.
FT DISULFID 836 849 BY SIMILARITY.
FT DISULFID 855 865 BY SIMILARITY.
FT DISULFID 860 874 BY SIMILARITY.
FT DISULFID 876 889 BY SIMILARITY.
FT DISULFID 958 970 BY SIMILARITY.
FT DISULFID 965 979 BY SIMILARITY.
FT DISULFID 981 994 BY SIMILARITY.
FT DISULFID 1076 1088 BY SIMILARITY.
FT DISULFID 1083 1097 BY SIMILARITY.
FT DISULFID 1099 1112 BY SIMILARITY.
FT DISULFID 1118 1130 BY SIMILARITY.
FT DISULFID 1125 1139 BY SIMILARITY.
FT DISULFID 1141 1155 BY SIMILARITY.
FT DISULFID 1161 1173 BY SIMILARITY.
FT DISULFID 1168 1182 BY SIMILARITY.
FT DISULFID 1184 1197 BY SIMILARITY.
FT DISULFID 1203 1215 BY SIMILARITY.
FT DISULFID 1210 1224 BY SIMILARITY.
FT DISULFID 1226 1239 BY SIMILARITY.
FT DISULFID 1245 1256 BY SIMILARITY.
FT DISULFID 1252 1265 BY SIMILARITY.
FT DISULFID 1267 1280 BY SIMILARITY.
FT DISULFID 1286 1298 BY SIMILARITY.
FT DISULFID 1293 1307 BY SIMILARITY.
FT DISULFID 1309 1322 BY SIMILARITY.
FT DISULFID 1328 1340 BY SIMILARITY.
FT DISULFID 1335 1349 BY SIMILARITY.
FT DISULFID 1351 1364 BY SIMILARITY.
FT DISULFID 1370 1383 BY SIMILARITY.
FT DISULFID 1377 1392 BY SIMILARITY.
FT DISULFID 1394 1405 BY SIMILARITY.
FT DISULFID 1411 1424 BY SIMILARITY.
FT DISULFID 1418 1433 BY SIMILARITY.
FT DISULFID 1435 1446 BY SIMILARITY.
FT DISULFID 1452 1464 BY SIMILARITY.
FT DISULFID 1459 1473 BY SIMILARITY.
```

```
Query Match 11.9%: Score 135; DB 1; Length 2911;
Best Local Similarity 23.9%: Pred. No. 0.00058;
Matches 60; Conservative 23; Mismatches 84; Indels 84; Gaps 15;
```

```
OY 6 LPQCAANKTCPTNMMWN---HICRC---LAQEDMFSSDADDBST----- 46
DB 2491 LDECSQSRPC--NYICKNTBGSYQSCPRGVYLOEDGKTKDCEQTKOHNCQFLCVN 2548
```

```
OY 47 -----DGF-----HDICGPNKELDEE-----TCQ-----CVCRAGL-----RP 74
DB 2549 TLGGFTCKRPPEFTQHTACIDNNECGSOPLLCGKGICQNTPGSFSCCQYGFSLDATG 2608
OY 75 ASCGPKKELDRN-SCQCYCKNKLFPSCGA-----NREFDENTCQCYCKRTCPRN 123
DB 2609 LNCEDVEDCEDGNHRCQHGQNTLGGYRCGCPQGYIOWYQMNOCVDENEC----- 2658
OY 124 OPLNKGK---ACECTESQKLL-KGKFFHQTCSCR-RECTNRQKACERGFYSYSE-- 176
DB 2659 ---NPACGSASCYNTLGSYKCAKCPGSPFDFQSSACHDVNCCSSKNPCNYGCSNTEGG 2715
OY 177 EVCRCVPSYWK 187
DB 2716 YLCGCPGYR 2726

RESULT 12
VWF_PIG STANDARD; PRT; 2482 AA.
ID VWF_PIG AC 028833:
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE VON WILLEBRAND FACTOR PRECURSOR (VWF) (FRAGMENT);
GN F8VWF OR VWF.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Seaman W.T., Read M.S., Bellinger D.A., Nichols T.C.;
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
RX [2]
RX SEQUENCE OF 397-553 FROM N.A.
RX MEDLINE=93356762; PubMed=8352759;
RA Laverne J.M., Piao Y.C., Ferreira V., Kerblirou-Nabias D.,
RA Bahnak B.R., Meyer D.;
RT "Primary structure of the factor VIII binding domain of human, porcine
RL and rabbit von Willebrand factor";
RL Biochem. Biophys. Res. Commun. 194:1019-1024(1993).
CC -!- FUNCTION: IMPORTANT IN THE MAINTENANCE OF HEMOSTASIS, IT
CC PARTICIPATES IN PLATELET-VESSEL WALL INTERACTIONS BY FORMING A
CC NONCOVALENT COMPLEX WITH COAGULATION FACTOR VIII AT THE SITE OF
CC VASCULAR INJURY (BY SIMILARITY).
CC -!- SUBUNIT: OLIGOMERIC (BY SIMILARITY).
CC -!- PTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR
CC INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 3 VWF-C DOMAINS.
CC -!- SIMILARITY: CONTAINS A C-TERMINAL CYSTEINE KNOT-LIKE DOMAIN (CTCK).
CC -!- SIMILARITY: SOME, TO SILKORM HEMOCYTIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF052036; AAC06229.1; -
DR EMBL; S64541; AAB27829.2; -
DR HSSP; P04275; IATZ
DR InterPro; IPR000359; -
DR InterPro; IPR001007; -
DR InterPro; IPR001846; -
DR InterPro; IPR002035; -
DR InterPro; IPR002919; -
DR Pfam; PF000007; Cys_knot; 1.
DR Pfam; PF000092; vwa; 3.
DR Pfam; PF000093; vwc; 3.
DR Pfam; PF000094; vwd; 3.
```

DR Pfam: PF01826; TIL; 3.
 DR PRINTS: PRO0365; ENDOTHELIN.
 DR PRINTS: PRO0453; VMFADOMAIN.
 DR PROSITE: PS00234; VMFA; 3.
 DR PROSITE: PS01208; VMFC; 3.
 DR PROSITE: PS01185; CTCK_1; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 KW Blood coagulation; Platelet; Glycoprotein; Extracellular matrix;
 KW Plasma; Endothelial cell; Repeat; Cell adhesion.
 FT NON_TER 1
 FT PROPEP <1 437 BY SIMILARITY.
 FT CHAIN 438 2482 VON WILLEBRAND FACTOR.
 FT 62 215 VMED 2.
 FT DOMAIN 438 461 AMINO-TERMINAL.
 FT DOMAIN 462 507 EL.
 FT DOMAIN 500 527 CX.
 FT DOMAIN 541 687 VMED 3.
 FT DOMAIN 947 1127 VMFA 1.
 FT DOMAIN 1167 1334 VMFA 2.
 FT DOMAIN 1360 1540 VMFA 3.
 FT DOMAIN 1619 1771 VMED 4.
 FT DOMAIN 1885 1930 E2.
 FT DOMAIN 1924 1997 VMFC 1.
 FT DOMAIN 2098 2164 VMFC 2.
 FT DOMAIN 2249 2319 VMFC 3.
 FT DOMAIN 2393 2481 CTCK.
 FT SITE 2176 2178 CELL ATTACHMENT SITE (POTENTIAL).
 FT DISULFID 441 482 BY SIMILARITY.
 FT DISULFID 450 478 BY SIMILARITY.
 FT DISULFID 484 495 BY SIMILARITY.
 FT DISULFID 541 670 BY SIMILARITY.
 FT DISULFID 563 705 BY SIMILARITY.
 FT DISULFID 572 667 BY SIMILARITY.
 FT DISULFID 588 595 BY SIMILARITY.
 FT DISULFID 734 758 BY SIMILARITY.
 FT DISULFID 745 785 BY SIMILARITY.
 FT DISULFID 763 765 BY SIMILARITY.
 FT DISULFID 827 839 BY SIMILARITY.
 FT DISULFID 823 843 BY SIMILARITY.
 FT DISULFID 800 804 BY SIMILARITY.
 FT DISULFID 870 873 BY SIMILARITY.
 FT DISULFID 908 911 BY SIMILARITY.
 FT DISULFID 942 1128 BY SIMILARITY.
 FT DISULFID 1338 1339 BY SIMILARITY.
 FT DISULFID 1355 1541 BY SIMILARITY.
 FT DISULFID 1548 1573 BY SIMILARITY.
 FT DISULFID 1568 1609 OR 1611 (BY SIMILARITY).
 FT DISULFID 1641 1792 BY SIMILARITY.
 FT DISULFID 1619 1754 BY SIMILARITY.
 FT DISULFID 1596 1757 BY SIMILARITY.
 FT DISULFID 1662 1670 BY SIMILARITY.
 FT DISULFID 2383 2443 BY SIMILARITY.
 FT DISULFID 2408 2457 BY SIMILARITY.
 FT DISULFID 2419 2473 BY SIMILARITY.
 FT DISULFID 2423 2475 BY SIMILARITY.
 FT DISULFID 2430 2480 BY SIMILARITY.
 FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 905 905 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1184 1184 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1243 1243 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1640 1640 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1892 1892 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1959 1959 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2026 2026 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2069 2069 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2215 2215 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2254 2254 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2459 2459 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 2482 AA; 272394 MW; D499B7DDEBACADD CRC64;

Query Match 11.9%; Score 134.5; DB 1; Length 2482;
 Best Local Similarity 25.1%; Pred. No. 0.00056;

Matches 60; Conservative 16; Mismatches 86; Indels 77; Gaps 15;
 QY 4 ATLLPQ-----QAAKTCPTNNMNNHICRC-LAEDF--MSSSPAGDSTGDFHD 51
 Db 1968 AAPAPGCGCEVARLROEAHQCC-----EYECVDLVSCDLPYPHCGGLQPTLTNG 2021
 QY 52 ICGPNKEIDETTCOC---CGRAGLRPASCGRKE-----LDRNSCCVCNKLFPSSQ 100
 Db 2022 ECRPN-----FTCACKRECPGRPLP-SCPPIHTPALRKTCCDEECACNCVNTLSCP 2075
 QY 101 CGANREDEMTCCQCKKRC-----PRNPPLNPGKACCECTESPOKCLKK 147
 Db 2076 LCVLASTVYNDGCT--TTTCLDPKVCVHRGTVPYGVGFEEGDCVCTCD-----LEDA 2128
 QY 148 KFHHCQSCYRRPCTNRQACRGPESY---SEHYC-RCVP-----SYMK 187
 Db 2129 VMGLRVAQCAQKPC---EDSCRPGFTYVYHBEBCGCKLPSACKVYIGSFGRDVSVMK 2184
 RESULT 13
 EMRL_MOUSE
 ID EMRL_MOUSE STANDARD; PRT: 931 AA.
 AC Q61549;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CELL SURFACE GLYCOPROTEIN EMRL PRECURSOR (EMRL HORMONE RECEPTOR)
 GN (CELL SURFACE GLYCOPROTEIN F4/80).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BAUB/C; TISSUE=Peritoneal cavity;
 RX MEDLINE=96132946; PubMed=8550607;
 RA McKnight A.J., Macfarlane A.J., Dri P., Turley L., Willis A.C.,
 Gordon S.;
 RT "Molecular cloning of F4/80, a murine macrophage-restricted cell
 RT surface glycoprotein with homology to the G-protein-linked
 RT transmembrane 7 hormone receptor family.";
 RL J. Biol. Chem. 271:486-489(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97312684; PubMed=9169125;
 RA Lin H.H., Stubbs L.J., Mucenski M.L.;
 RT "Identification and characterization of a seven transmembrane hormone
 RT receptor using differential display.";
 RL Genomics 41:301-308(1997).
 CC -!- FUNCTION: PROBABLY INVOLVED IN CELL ADHESION WITHIN TISSUES
 CC AND RECEPTOR SIGNALING.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: IN MACROPHAGES; BUT ABSENT FROM THOSE WHICH
 CC ARE LOCALIZED WITHIN T-CELL AREAS OF LYMPH NODES AND SPLEEN.
 CC LOW LEVEL OF EXPRESSION ON BLOOD MONOCYTES.
 CC -!- SIMILARITY: CONTRAINS 7 EGF-LIKE DOMAINS.
 CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 DR EMBL: X93328; CA63720.1; -;
 DR EMBL: U66888; AAC53184.1; -;
 DR HSRF: P07204; IFGD.
 DR GCRDB: GCR_1309; -;
 DR MGD: MGI:106912; Emcl.
 DR InterPro: IPR000152; -;

DR InterPro: IPR000203; -
 DR InterPro: IPR000561; -
 DR InterPro: IPR000832; -
 DR InterPro: IPR001740; -
 DR InterPro: IPR001881; -
 DR Pfam: PF00002; 7cm_2; 1.
 DR Pfam: PF00008; EGF; 7.
 DR PRINTS: PR01128; EMR1HOMONER.
 DR PROSITE: PS00650; G_PROTEIN_RECIP_F2_2; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 6.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS01187; EGF_CA; 5.
 DR G-protein coupled receptor; Transmembrane; Receptor; Glycoprotein;
 KW EGF-like domain; Repeat; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 931
 FT DOMAIN 28 931
 FT TRANSMEM 645 644
 FT TRANSMEM 673 672
 FT TRANSMEM 680 701
 FT TRANSMEM 702 711
 FT TRANSMEM 712 735
 FT TRANSMEM 736 754
 FT TRANSMEM 755 776
 FT TRANSMEM 777 792
 FT TRANSMEM 793 821
 FT TRANSMEM 822 839
 FT TRANSMEM 840 859
 FT TRANSMEM 860 874
 FT TRANSMEM 875 897
 FT TRANSMEM 898 931
 FT DOMAIN 81 32
 FT DOMAIN 133 132
 FT DOMAIN 173 221
 FT DOMAIN 222 371
 FT DOMAIN 319 367
 FT DOMAIN 506 508
 FT SITE 36 48
 FT DISULFID 42 57
 FT DISULFID 59 79
 FT DISULFID 85 98
 FT DISULFID 92 107
 FT DISULFID 109 131
 FT DISULFID 137 149
 FT DISULFID 143 158
 FT DISULFID 160 171
 FT DISULFID 177 189
 FT DISULFID 183 198
 FT DISULFID 200 220
 FT DISULFID 226 239
 FT DISULFID 233 248
 FT DISULFID 250 270
 FT DISULFID 276 286
 FT DISULFID 297 317
 FT DISULFID 323 336
 FT DISULFID 330 345
 FT DISULFID 347 366
 FT CARBOHYD 148 167
 FT CARBOHYD 167 167
 FT CARBOHYD 229 229
 FT CARBOHYD 269 269
 FT CARBOHYD 283 283
 FT CARBOHYD 405 405
 FT CARBOHYD 417 417
 FT CARBOHYD 474 474
 FT CARBOHYD 498 498
 FT CARBOHYD 706 706
 SEQUENCE 931 AA: 102129 MW: 52963A67EBB76B5 CRC64;

Query Match 11.5%; Score 130; DB 1; Length 931;
 Best Local Similarity 21.6%; Pred. No. 0.00062;
 Matches 65; Conservative 18; Mismatches 76; Indels 142; Gaps 16;
 QY 8 OCOAANKTCPTNYMMNNH-----CRCLAQEDPMFSSDAGD-----STDGF-HDI----- 52
 DB 84 ECLQSDSPCGPNSVCTNITLGRKACSLRG-----FSSFGKDWILGSLDNLFCADVDECLT 139
 QY 53 -----CGPNKEL-----DEETC-----OCV 67
 DB 140 IGICPKYNSGNSVSGSCCTCPGFLVNGSLCEDEDECVTROVCEHATCNTLASYCT 199
 QY 68 CAGLRPASCGRH-----KELDRNSGQC-----VCKNKLPPSGCAGAREDEMTIC 112
 DB 200 CNSGLESSCGGPMFGGLDESCEDVDECSNLTLCGPTFCITWTLSYSCSCAGSFLPTF 259
 QY 113 QCV-----CKRCPRN-----QPLNPKACAC----- 133
 DB 260 QILGHPADGNCCTDIDCEDTCTPLNSSCTWTIGSYCTCHPGFASNGQLNFKLEVTGCD 319
 QY 134 --ECIESPGKCLLKGRKHHQSCYRRPCTRQKACRPGFSYSEVGRVPSYKRRPM 191
 DB 320 IDECTQDPLQGLNS-----VCTN-----VPG-SY-----ICGCLPDPQMDPEG 357
 QY 192 S 192
 DB 358 S 358

RESULT 14
 TENA_HUMAN
 AC P24821; Q15567; Q14583; PRT; 2201 AA.
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE TENASCIN PRECURSOR (TN) (HEXABRACHION) (CYTOTACTIN) (NEURONECTIN)
 DE (GEM) (JI) (MIOENDINIUS ANTIGEN) (GLIOMA-ASSOCIATED-EXTRACELLULAR
 DE MATRIX ANTIGEN) (GP 150-225) (TENASCIN-C) (TN-C).
 GN HXB
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-32.
 RC TISSUE=Fetal brain, and Melanoma;
 RX MEDLINE=91187670; PubMed=1707164;
 RA Sirri A., Carmemolla B., Saginatti M., Lepri A., Casari G.,
 RA Baralle F., Zardi L.;
 RT "Human tenascin: primary structure, pre-mRNA splicing patterns and
 RT localization of the epitopes recognized by two monoclonal
 RT antibodies.";
 RL Nucleic Acids Res. 19:525-531(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95155442; PubMed=7531707;
 RA Cherzi R., Carmemolla B., Sirri A., Ponassi M., Balza E., Zardi L.;
 RT "Human tenascin gene. Structure of the 5'-region, identification, and
 RT characterization of the transcription regulatory sequences.";
 RL J. Biol. Chem. 270:3429-3434(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91131572; PubMed=1704365;
 RA Nies D.E., Hemesath T.J., Kim J.H., Gulcher J.R., Stefansson K.;
 RT "The complete cDNA sequence of human hexabrachion (tenascin). A
 RT multidomain protein containing unique epidermal growth factor
 RT repeats.";
 RL J. Biol. Chem. 266:2818-2823(1991).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92052108; PubMed=1719530;

```

RA  Gulcher J.R., Nies D.E., Alexakos M.J., Ravikant N.A., Sturgill M.E.,
RA  Marton L.S., Stefansson K.;
RA  "Structure of the human hexabrachion (tenascin) gene.";
RA  Proc. Natl. Acad. Sci. U.S.A. 88:9438-9442(1991).
RN  [5]
RN  SEQUENCE OF 431-2055 FROM N.A. (ISOFORM P31).
RP  TISSUE=Glolioblastoma;
RX  MEDLINE=89160821; PubMed=2466295;
RA  Gulcher J.R., Nies D.E., Marton L.S., Stefansson K.;
RA  "An alternatively spliced region of the human hexabrachion contains a
RA  repeat of potential N-glycosylation sites.";
RA  Proc. Natl. Acad. Sci. U.S.A. 86:1588-1592(1989).
RN  [6]
RN  X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF FIBRONECTIN TYPE-III 3.
RX  MEDLINE=93068293; PubMed=1279805;
RA  Leahy D.J., Hendrickson W.A., Aukhil I., Erickson H.P.;
RA  "Structure of a fibronectin type III domain from tenascin phased by
RA  MAD analysis of the selenomethionyl protein.";
RA  Science 258:987-991(1992).
CC  -I- FUNCTION: SAM (SUBSTRATE-ADHESION MOLECULE) THAT APPEARS TO
CC  INITIATE CELL MIGRATION. MAY PLAY A ROLE IN SUPPORTING THE GROWTH
CC  OF EPITHELIAL TUMORS.
CC  -I- SUBUNIT: HEXAMERIC. A HOMOTRIMER MAY BE FORMED IN THE TRIPLE
CC  COILED-COIL REGION AND MAY BE STABILIZED BY DISULFIDE RINGS AT
CC  BOTH ENDS. TWO OF SUCH HALF-HEXABRACHIONS MAY BE DISULFIDE LINKED
CC  WITHIN THE CENTRAL GLOBULE.
CC  -I- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
CC  -I- ALTERNATIVE PRODUCTS: FOUR VARIANTS ARE PRODUCED FROM A SINGLE
CC  GENE IN A TISSUE- AND TIME-SPECIFIC MANNER DURING DEVELOPMENT.
CC  -I- INDUCTION: BY TGF-BETA.
CC  -I- SIMILARITY: CONTAINS 15 EGF-LIKE DOMAINS.
CC  -I- SIMILARITY: CONTAINS 15 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC  -I- SIMILARITY: CONTAINS 1 FIBRINOGEN-LIKE DOMAIN.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL, X56160; CA39628.1; -
DR  EMBL, X78565; CA55309.1; -
DR  EMBL, M55618; AAA88083.1; -
DR  EMBL, M24630; AAA52703.1; -
DR  PIR, S14015; S14015
DR  PDB, ITEM: 31-OCT-93.
DR  MIM, 187380; -
DR  InterPro: IPR000561; -
DR  InterPro: IPR001777; -
DR  InterPro: IPR002181; -
DR  Pfam, PF000008; EGF_14.
DR  Pfam, PF00147; fibrinogen_C; 1.
DR  Pfam, PF00041; fn3; 15.
DR  PROSITE, PS00022; EGF_1; 14.
DR  PROSITE, PS01186; EGF_2; 15.
KW  Glycoprotein; Cell adhesion; Repeat; EGF-like domain; Coiled coil;
KW  Extracellular matrix; Signal; Alternative splicing; 3D-structure.
FT  SIGNAL 1 22
FT CHAIN 23 2201
FT DOMAIN 118 145 TENASCIN.
FT DOMAIN 174 186 COILED COIL (POTENTIAL).
FT DOMAIN 217 248 EGF-LIKE 1 (INCOMPLETE).
FT DOMAIN 248 280 EGF-LIKE 2.
FT DOMAIN 280 311 EGF-LIKE 3.
FT DOMAIN 311 342 EGF-LIKE 4.
FT DOMAIN 342 373 EGF-LIKE 5.
FT DOMAIN 373 404 EGF-LIKE 6.
FT DOMAIN 404 435 EGF-LIKE 7.
FT DOMAIN 435 466 EGF-LIKE 8.
FT DOMAIN 466 497 EGF-LIKE 9.
FT DOMAIN 497 EGF-LIKE 10.
FT DOMAIN 497 EGF-LIKE 11.

```

FT	DOMAIN	497	528	EGF-LIKE 12.
FT	DOMAIN	528	559	EGF-LIKE 13.
FT	DOMAIN	559	590	EGF-LIKE 14.
FT	DOMAIN	590	621	EGF-LIKE 15.
FT	DOMAIN	621	710	FIBRONECTIN TYPE-III 1.
FT	DOMAIN	711	801	FIBRONECTIN TYPE-III 2.
FT	DOMAIN	802	891	FIBRONECTIN TYPE-III 3.
FT	DOMAIN	892	983	FIBRONECTIN TYPE-III 4.
FT	DOMAIN	984	1071	FIBRONECTIN TYPE-III 5.
FT	DOMAIN	1072	1162	FIBRONECTIN TYPE-III 6.
FT	DOMAIN	1163	1253	FIBRONECTIN TYPE-III 7.
FT	DOMAIN	1254	1345	FIBRONECTIN TYPE-III 8.
FT	DOMAIN	1345	1435	FIBRONECTIN TYPE-III 9.
FT	DOMAIN	1436	1526	FIBRONECTIN TYPE-III 10.
FT	DOMAIN	1527	1617	FIBRONECTIN TYPE-III 11.
FT	DOMAIN	1618	1708	FIBRONECTIN TYPE-III 12.
FT	DOMAIN	1709	1797	FIBRONECTIN TYPE-III 13.
FT	DOMAIN	1798	1885	FIBRONECTIN TYPE-III 14.
FT	DOMAIN	1886	1973	FIBRONECTIN TYPE-III 15.
FT	DOMAIN	1982	2201	FIBRINGEN BETA/GAMMA.
FT	DISULFID	190	200	BY SIMILARITY.
FT	DISULFID	194	205	BY SIMILARITY.
FT	DISULFID	207	216	BY SIMILARITY.
FT	DISULFID	221	231	BY SIMILARITY.
FT	DISULFID	225	236	BY SIMILARITY.
FT	DISULFID	238	247	BY SIMILARITY.
FT	DISULFID	252	263	BY SIMILARITY.
FT	DISULFID	256	268	BY SIMILARITY.
FT	DISULFID	270	279	BY SIMILARITY.
FT	DISULFID	284	294	BY SIMILARITY.
FT	DISULFID	288	299	BY SIMILARITY.
FT	DISULFID	301	310	BY SIMILARITY.
FT	DISULFID	315	325	BY SIMILARITY.
FT	DISULFID	319	330	BY SIMILARITY.
FT	DISULFID	332	341	BY SIMILARITY.
FT	DISULFID	346	356	BY SIMILARITY.
FT	DISULFID	350	361	BY SIMILARITY.
FT	DISULFID	363	372	BY SIMILARITY.
FT	DISULFID	377	387	BY SIMILARITY.
FT	DISULFID	391	392	BY SIMILARITY.
FT	DISULFID	394	403	BY SIMILARITY.
FT	DISULFID	408	418	BY SIMILARITY.
FT	DISULFID	412	423	BY SIMILARITY.
FT	DISULFID	425	434	BY SIMILARITY.
FT	DISULFID	439	449	BY SIMILARITY.
FT	DISULFID	443	454	BY SIMILARITY.
FT	DISULFID	456	465	BY SIMILARITY.
FT	DISULFID	470	480	BY SIMILARITY.
FT	DISULFID	474	485	BY SIMILARITY.
FT	DISULFID	487	496	BY SIMILARITY.
FT	DISULFID	501	511	BY SIMILARITY.
FT	DISULFID	505	516	BY SIMILARITY.
FT	DISULFID	518	527	BY SIMILARITY.
FT	DISULFID	532	542	BY SIMILARITY.
FT	DISULFID	536	547	BY SIMILARITY.
FT	DISULFID	549	558	BY SIMILARITY.
FT	DISULFID	563	573	BY SIMILARITY.
FT	DISULFID	567	578	BY SIMILARITY.
FT	DISULFID	580	589	BY SIMILARITY.
FT	DISULFID	594	604	BY SIMILARITY.
FT	DISULFID	598	609	BY SIMILARITY.
FT	DISULFID	611	620	BY SIMILARITY.
FT	CARBOHYD	38	38	N-LINKED (GLCNAC. .)
FT	CARBOHYD	166	166	N-LINKED (GLCNAC. .)
FT	CARBOHYD	184	184	N-LINKED (GLCNAC. .)
FT	CARBOHYD	327	327	N-LINKED (GLCNAC. .)
FT	CARBOHYD	788	788	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1018	1018	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1034	1034	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1079	1079	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1093	1093	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1119	1119	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1184	1184	N-LINKED (GLCNAC. .)

```
FT CARBOHYD 1210 1210 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1261 1261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1275 1275 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1301 1301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1366 1366 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1392 1392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1445 1445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1455 1455 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1485 1485 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1534 1534 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1809 1809 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2162 2162 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 1072 1435 MISSING (IN ISOFORM HT-5 AND ISOFORM
Query Match 11.5%; Score 130; DB 1; Length 2201;
Best Local Similarity 23.6%; Pred. No. 0.0012;
Matches 54; Conservative 20; Mismatches 83; Indels 72; Gaps 15;
QY 13 NKTCPNYMNNHIC---RCLAEDEFSSDAGDSTDFHDICGPNKEIDERTC---QC 66
DB 280 NKPLCLNNCYNRRCVENCVCDEGF---TGEDCSE---LIC-PNDCFDRCINGTC 330
QY 67 VCRAGLRPASCG---PHKEIDRNSC---QCV-----CKNKLFPPQCCANREFDEN 110
DB 331 YCEEGFTGEDCGKPTCPHACHTGTGCEGQCVCEGFAGLDCSEKRCPADCHNRGCVDG 390
QY 111 TCQC-----VCKRTCPRNQPLNPKGKAC-----ECTE--SPQKCLKGK-- 147
DB 391 RCECDGFTGADGELKRCNGCSGHRGCVNGQCVCEGTYGECSSLRCPNCHSGRCV 450
QY 148 -----KFHHOTCS-----CYRR-PCYNRQKACERGFYSYSEVCR 180
DB 451 ESKVCCEGQFGKGYDCSDMSCPNDCHQHRGCVNGMCVCDG--YTGEDCR 497
RESULT 15
PCK5_BRACL STANDARD; PRT: 1696 AA.
AC Q9NJ15; Q9NJ16; Q9NJ14;
DT 01-OCT-2000 (Rel. 40; Created)
DT 01-OCT-2000 (Rel. 40; Last sequence update)
DT 01-OCT-2000 (Rel. 40; Last annotation update)
DE PROPROTEIN CONVERTASE SUBTILISIN/KEXIN TYPE 5 PRECURSOR (EC 3.4.21.-)
(GN) (PROPROTEIN CONVERTASE PC6-LIKE) (APC6).
OS Branchiostoma californiensis (California lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OX NCBI_TaxID=7738;
RN 111
RP SEQUENCE FROM N.A. (ISOFORM A; ISOFORM B AND ISOFORM C).
RX MEDLINE=20175281; PubMed=10708868;
RA Oliva A.A. Jr., Chan S.J., Steiner D.F.;
RT "Evolution of the prothormone convertases: identification of a
RT homologue of PC6 in the protochordate amphioxus."
RL Biochim. Biophys. Acta 1477:338-348(2000).
CC -1- FUNCTION: LIKELY TO REPRESENT A WIDESPREAD SECRETORY PATHWAY CAPABLE
CC WITHIN THE CONSTITUTIVE AND REGULATED SECRETORY PATHWAY. CAPABLE
CC OF CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR
CC PROPEPTIDES BY CLEAVAGE OF ARG-XAA-YAA-ARG-1-2AA BONDS, WHERE XAA
CC CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
CC -1- SUBCELLULAR LOCATION: ISOFORM A AND ISOFORM C ARE SECRETED.
CC ISOFORM B IS A TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A, B (SHOWN HERE) AND C; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE
CC ASSISTING THE FOLDING OF THE ZMOGEN WITHIN THE ENDOPLASMIC
CC RETICULUM.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC -1- SIMILARITY: CONTAINS 1 HOMO B/P DOMAIN.
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AF184615; AAF26300.1; -
CC DR EMBL; AF184616; AAF26301.1; -
CC DR EMBL; AF184617; AAF26302.1; -
CC DR PROSITE; PS00136; SUBTILASIN; FALSE_NEG.
CC DR PROSITE; PS00137; SUBTILASIN; HIS; 1.
CC DR PROSITE; PS00138; SUBTILASIN; SER; 1.
CC KW Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;
CC Cleavage on pair of basic residues; Repeat; Alternative splicing;
CC Transmembrane.
CC FT SIGNAL 1 25
CC FT PROPEP 26 110
CC FT CHAIN 111 1696
CC FT DOMAIN 111 1618
CC FT TRANSMEM 1619 1639
CC FT DOMAIN 1640 1696
CC FT DOMAIN 111 488
CC FT DOMAIN 496 637
CC FT DOMAIN 664 1649
CC FT SITE 110 111
CC FT ACT_SITE 192 192
CC FT ACT_SITE 233 233
CC FT ACT_SITE 407 407
CC FT CARBOHYD 246 246
CC FT CARBOHYD 529 529
CC FT CARBOHYD 885 885
CC FT VARSPLIC 1259 1323
CC FT PREVSIALALGHLRLYSITDVPQSNPSPTVLGADRRL
CC FT TTATSAAGRA (IN ISOFORM C).
CC FT MISSING (IN ISOFORM C).
CC FT CHPTCKECSDEYDPTCAKNGGFLITDASSGACPPQFL
CC FT HHGDCDSCHRECKTC -> IARCVDRRRSCDLYRNFC
CC FT VRRYFVRRCCGCTKLYMEDRPMRRSSQPTQGRN (IN
CC ISOFORM A).
CC FT MISSING (IN ISOFORM A).
CC FT VARSPLIC 1344 1696
CC FT SEQUENCE 1696 AA; 188410 MW; 281CBEL784257CBD CRC64;
Query Match 11.4%; Score 128.5; DB 1; Length 1696;
Best Local Similarity 19.3%; Pred. No. 0.0013;
Matches 48; Conservative 29; Mismatches 63; Indels 109; Gaps 12;
QY 11 AANKTCPTNYMNNHICRL---AEDDEFSSDAG--DD-----STDFHDICGR- 56
DB 1146 SCRTCPNAGFTGNASHHECVSSCEODDYSETGRCEDCPYNCACNDGDCACAPTY 1205
QY 57 -----KEIDERTCQVCR-----A 70
DB 1206 IYVDRCRREFTCEGEYQDRPRDIAELSCRCRCHOSCTKCSPTDSDCKGDTILDR 1265
QY 71 GLRPASCGPHKEIDR---NSQCCCKNKLFPSSCGANREFDNTCC-----V 115
DB 1266 GECITSQGEYMDRBRKCKKCHPTCK-----EC--SDEYDPTCAKNGGFLITDASS 1317
QY 116 CKRTCPRNQPLNPKG-----ACECTESPQKCLKGKFFHHOTCSYRRPCTNRQKACEP 170
DB 1318 CEAGCPPOQLFHHGDCDSCHRECKTCGP-----HHQNC-----LSCOP 1356
QY 171 GFYSYSEEV 179
DB 1357 GSYLNDQOC 1365
```

Search completed: October 17, 2001, 14:51:45
Job time: 471 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:50:53 ; Search time 128.06 Seconds
(without alignments)
198.365 Million cell updates/sec

Title: US-09-427-657-2_COPY_228_419

Perfect score: 1130
Sequence: 1 SLPALPQCAANKTCPTNY.....SYSEVCRCPVSYWKRPQMS 192

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_16: *
1: sp.archaea: *
2: sp.bacteria: *
3: sp.fungi: *
4: sp.human: *
5: sp.invertebrate: *
6: sp.mammal: *
7: sp.mmc: *
8: sp.organelle: *
9: sp.phage: *
10: sp.plant: *
11: sp.prodent: *
12: sp.unclassified: *
13: sp.vertebrate: *
14: sp.virus: *

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	991	87.7	420	6 Q9XS50	Q9XS50 bos taurus
2	839	74.2	418	13 057352	057352 coturnix co
3	257	22.7	126	11 035757	035757 rattus norv
4	245.5	21.7	1704	5 Q94446	Q94446 chironomus
5	235.5	20.8	1698	5 Q94438	Q94438 chironomus
6	234	20.7	354	11 Q93915	Q93915 homo sapien
7	225	19.9	358	11 P97946	P97946 mus musculu
8	199	17.6	326	11 035251	035251 rattus norv
9	171.5	15.2	704	3 074567	074567 trichoderma
10	151	13.4	1810	13 Q90824	Q90824 gallus gall
11	148	13.1	800	5 Q9VB78	Q9VB78 drosophila
12	148	13.1	1532	13 Q90994	Q90994 gallus gall
13	148	13.1	1714	13 Q90995	Q90995 gallus gall
14	145	12.8	543	5 Q9VJ05	Q9VJ05 drosophila
15	145	12.8	620	5 Q9NKB8	Q9NKB8 drosophila
16	145	12.8	830	4 Q43701	Q43701 homo sapien
17	145	12.8	830	4 Q4162	Q4162 homo sapien
18	144	12.7	663	5 Q4247	Q4247 hemientrot
19	139.5	12.3	1203	11 Q06008	Q06008 mus musculu

20	139.5	12.3	2470	11 Q35516	Q35516 mus musculu
21	139.5	12.3	2471	4 Q04721	Q04721 homo sapien
22	139.5	12.3	2471	4 Q9H240	Q9H240 homo sapien
23	139.5	12.3	2471	11 Q9QW50	Q9QW50 rattus sp.
24	139.5	12.3	2475	5 Q9GP97	Q9GP97 caenorhabdi
25	138.5	12.3	765	5 Q9NL50	Q9NL50 sarcophaga
26	138	12.2	648	5 Q9NKD7	Q9NKD7 drosophila
27	138	12.2	701	5 Q9VJ04	Q9VJ04 drosophila
28	136.5	12.1	122	6 Q9GLX1	Q9GLX1 bos taurus
29	136	12.0	777	5 Q24550	Q24550 drosophila
30	136	12.0	777	5 Q9VYPI	Q9VYPI drosophila
31	136	12.0	2653	5 Q25253	Q25253 lucilia cup
32	135.5	12.0	1187	2 Q49549	Q49549 mycoplasma
33	135.5	12.0	1511	5 Q9YB21	Q9YB21 drosophila
34	135.5	12.0	4006	11 Q35452	Q35452 mus musculu
35	135.5	12.0	4114	11 Q54796	Q54796 mus musculu
36	135	11.9	778	13 Q9IBG4	Q9IBG4 xenopus lae
37	135	11.9	1587	4 Q00508	Q00508 homo sapien
38	134.5	11.9	999	4 Q9N036	Q9N036 homo sapien
39	134	11.9	642	13 P79941	P79941 xenopus lae
40	134	11.9	1218	11 Q9QXX0	Q9QXX0 mus musculu
41	134	11.9	1219	11 Q63722	Q63722 rattus norv
42	134	11.9	1664	5 Q9TVQ2	Q9TVQ2 caenorhabdi
43	134	11.9	2352	5 Q61240	Q61240 halocynthia
44	134	11.9	2906	11 Q9WU9	Q9WU9 rattus norv
45	133	11.8	2019	11 Q64706	Q64706 mus musculu

ALIGNMENTS

RESULT	ID	Q9XS50	PRELIMINARY	PRT	420 AA
Q9XS50	Q9XS50	Q9XS50	Q9XS50	Q9XS50	Q9XS50
DR	01-NOV-1999	(TREMBLrel. 12, Created)			
DT	01-NOV-1999	(TREMBLrel. 12, Last sequence update)			
DT	01-MAR-2001	(TREMBLrel. 16, Last annotation update)			
DE	VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR.				
OS	Bos taurus (Bovine).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;				
OC	Bovidae; Bovinae; Bos.				
OX	NCBI_TaxID=9913;				
RM	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=HEART;				
RA	Liu X., Yonekura H., Yamagishi S., Yamamoto Y., Yamamoto H.;				
RT	"Structure and expression of bovine VEGF family."				
RL	Submitted (May-1997) to the EMBL/Genbank/DBJ databases.				
DR	EMBL; AB004275; BAA77687.1; -.				
DR	HSSP; P15692; IVP.				
DR	InterPro; IPR000072; -.				
DR	Pfam; PF00341; PDGF_1.				
DR	PROSITE; PS00249; PDGF_1; 1.				
DR	PROSITE; PS0278; PDGF_2; 1.				
DR	SMART; SM00141; PDGF_1.				
FM	Signal.				
KW	SIGNAL.				
FT	CHAIN	1	20	POTENTIAL.	
FT	SEQUENCE	420 AA;	46681 MM;	58BA84317A3CE2D CXC64;	
SO					
QY	1	SLPALPQCAANKTCPTNYMMNNHICRLAEDPMFSSDAGDSDTDFGHDICGKNEID	60		
DB	229	SLPALPQCAANKTCPADYINNNHVCRLAHDHIFPSADGDSADGFHDICGKNEID	288		
OY	61	ETTCQCVCRAGLRPAISCGPHKELDRNSCQVCYKNTLPSGCGANPEPNTCCQVCKRRC	120		

DB 289 EETCCGVCAGLQASSCGPHKELRDSCOCVCNKLFPSSCGANREDFENFCCICKKTC 348
QY 121 PPNOPINPKACCECTESPCKLKKRHHOTCSYRPPCNROKACBPGRSYSEVCR 180
DB 349 PPNOPINPKACCECTESPCKLKKRHHOTCSYRPPCNROKACBPGRSYSEVCR 408
QY 181 CVPSTWKRPPOMS 192
DB 409 CVPSTWKRPPHVN 420
RESULT 2
057352 PRELIMINARY; PRT; 418 AA.
AC 057352;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR.
GN VEGF-C.
OS Colurnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98167900; PubMed=9435294;
RA Eichmann A., Corbel C., Jaffredo T., Breant V., Joukov V., Kumar V.,
RA Allitalo K., Le Douarin N.M.;
RT "Avian VEGF-C: cloning, embryonic expression pattern and stimulation
RT of the differentiation of VEGFR2-expressing endothelial cell
RT precursors.";
RT Development 125:743-752(1998).
DR EMBL: Y15837; CAA75799.1; -.
DR HSSP: P15692; 1VPF.
DR InterPro: IPR000072; -.
DR InterPro: IPR002400; -.
DR Pfam: PF00341; PDGF_1.
DR PRINTS: PR00438; GFCYSKNOT.
DR PRODOM: PD001629; -. 1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS00278; PDGF_2; 1.
DR SMART: SM00141; PDGF_1.
DR Signal.
KM SIGNAL.
FT CHAIN 111 418 .POTENTIAL.
FT SIGNAL 111 418 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
SQ SEQUENCE 418 AA; 46839 MW; 099BFC79151BF2B CRC64;

Query Match 74.2%; Score 839; DB 13; Length 418;
Best Local Similarity 73.1%; Pred. No. 1.4e-82;
Matches 141; Conservative 17; Mismatches 33; Indels 2; Gaps 2;
QY 1 SLPATLPQCOAANKTCPTNYMNNHICRCLAODEFMSSDAD-DSTGCFHICGPNKL 59
DB 227 SLPATLPQCOAANKTCPTNYMNNHICRCLAODEFMSSDAD-DSTGCFHICGPNKL 285
QY 60 DEFTQCVCRAGLRPASCPHKELRDSCOCVCNKLFPSSCGANREDFENFCCICKKTC 119
DB 286 DEFTQCVCRAGLRPASCPHKELRDSCOCVCNKLFPSSCGANREDFENFCCICKKTC 345
QY 120 CPNQLPMPGKACCECTESPCKLKKRHHOTCSYRPPCNROKACBPGRSYSEVCR 179
DB 346 CPNQLPMPGKACCECTESPCKLKKRHHOTCSYRPPCNROKACBPGRSYSEVCR 405
QY 180 CVPSTWKRPPOMS 192
DB 406 CVPSTWKRPPHVN 420

RESULT 3
035757

ID 035757 PRELIMINARY; PRT; 126 AA.
AC 035757;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR C (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LUNG;
RA Mandriota S.J., Pepper M.S.;
RT Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF010302; AAB63248.1; -.
DR HSSP: P15692; 2VPF.
DR InterPro: IPR000072; -.
DR Pfam: PF00341; PDGF_1.
DR PROSITE: PS00278; PDGF_2; 1.
DR SMART: SM00141; PDGF_1.
FT NON_TER 1 126
FT NON_TER 126 126
SQ SEQUENCE 126 AA; 13977 MW; 8F365AFBC4E037B0 CRC64;

Query Match 22.7%; Score 257; DB 11; Length 126;
Best Local Similarity 74.5%; Pred. No. 2.4e-20;
Matches 41; Conservative 10; Mismatches 4; Indels 0; Gaps 0;

QY 1 SLPATLPQCOAANKTCPTNYMNNHICRCLAODEFMSSDAD-DSTGCFHICGPNKL 55
DB 72 SLPATLPQCOAANKTCPTNYMNNHICRCLAODEFMSSDAD-DSTGCFHICGPNKL 126

RESULT 4
094446 PRELIMINARY; PRT; 1704 AA.
AC 094446;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE 220 KDA SILK PROTEIN.
GN SP220.
OS Chironomus thummi thummi (Midge).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomidae; Chironomidae; Chironominae; Chironomus.
OX NCBI_TaxID=7155;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPRAGUE-DAWLEY; TISSUE=LUNG;
RA Case S.T., Cox C., Bell W.C., Hoffman R.T., Martin J., Hamilton R.;
RT Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U54641; AAA9604.1; -.
DR InterPro: IPR000561; -.
DR InterPro: IPR000853; -.
DR PRINTS: PR00876; MTNEMATODE.
DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
SQ SEQUENCE 1704 AA; 185745 MW; 3A3F20247C8F1E28 CRC64;

Query Match 21.7%; Score 245.5; DB 5; Length 1704;
Best Local Similarity 27.6%; Pred. No. 3.8e-18;
Matches 56; Conservative 31; Mismatches 75; Indels 41; Gaps 9;

QY 3 PATLPQCOAANKTCPTNYMNNHICRCLAODEFMSSDAD-DSTGCFHICGPNKL 45
DB 1210 PATLPQCOAANKTCPTNYMNNHICRCLAODEFMSSDAD-DSTGCFHICGPNKL 1265
QY 46 TDGFDICGPNKELEDFQCVCRAGLRPASCPHKELRDSCOCVCNKLFPSSCGAN 104

Db 1266 -----CG-NKMWNDKACECECKPCGPEAGCKGVOKWKNKTCACBOPPKATPASCIDK 1317
 OY 105 REFEDNTCOCVCKRT-----CPRNQPLNPGKCAECETESPQCKLLGKKPHHOTCSCYRR 159
 Db 1318 KSNWNPSCSCQCKSKMPGGGCPSPNOQNCETCCECS-GRTOC-PAGQSDSOTCOC-SC 1374
 OY 160 PCTNRKACEPGFSPSEYRCV 182
 Db 1375 PATGK---CTGAQFWCAKCKCKV 1394

RESULT 5
 OY4438 PRELIMINARY: PRT: 1698 AA.
 ID OY4438:
 AC OY4438:
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 DE 185 KDA SILK PROTEIN.
 GN SP185.
 OS Chironomus pallidivittatus (Midge).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
 OC Chironomidae; Chironomidae; Chironomidae; Chironomus.
 NCBI_TaxId=7151;
 [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Salivary Gland;
 RC Case S.T., Cox C., Bell W.C., Hoffman R.T., Martin J., Hamilton R.;
 RA Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 RL EMBL; 054640; AAA9803.1;
 DR InterPro: IPR000853;
 DR InterPro: IPR000853;
 DR PRINTS: PRO0876; MTNEMATODE.
 DR PROSITE: PS00022; EGF_1; UNKNOMN_1.
 SQ SEQUENCE 1698 AA; 186164 MW; 82AF0838D9E2360E CRC64;

Query Match 20.8%; Score 235.5; DB 5; Length 1698;
 Best Local Similarity 25.7%; Pred. No. 4,5e-17;
 Matches 63; Conservative 32; Mismatches 79; Indels 71; Gaps 12;

OY 1 SUPATLP-----QC--QAAN-KTCPTNYMNNHICRLAQEDMFSSDAGDSDTDF 49
 Db 1027 AMPATIPPKKKNKEDKVCLECANVKTCEBPQRKCDNCKCICQ-----V 1071
 OY 50 HDICGPNKELDEETCC-----VCRAGLR-----PASCGRPKELDRNSCQVC 92
 Db 1072 NTKCSDBKQKFIKSCBEGDPTOTCKNGFRWSNLECGCLDEKCKGQVDFDKNTCCCK 1131
 OY 93 KKNLFPSQCGAREPENTCQVC-----RTCPRNQPLNPGKACAC-TESPQCKLLG 146
 Db 1132 PKNKPEDTGNNGDFCPLDSCCKSPKANGCPGVQENNEEKQCECKRKRKCKP 1191
 OY 147 KKEFHOTCSC---YRPCTNRK-----ACEPFS-----YSEECRCV-PS 184
 Db 1192 QDMNNHLCCGCGPTAPTCOSNNQKYSNVSCGCKGRKREGRPKQIMCENTRCVCK 1251
 OY 185 YMKRP 189
 Db 1252 NMDKP 1256

RESULT 6
 OY43915 PRELIMINARY: PRT: 354 AA.
 ID OY43915:
 AC OY43915:
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
 DE GROWTH FACTOR FIGF.
 GN FIGF OR VEGF-D.
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxId=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE-98140120; PubMed-9479493;
 RX Rocchigliani M., Testi M., Luddi A., Orlandini M., Franco B.,
 RA Rossi E., Ballabio A., Zuffardi O., Oliviero S.;
 RT "Human FIGF: cloning, gene structure, and mapping to chromosome Xp22.1
 between the FIGA and the GPR genes";
 RL Genomics 47:207-216(1998).
 [2]
 RN SEQUENCE FROM N.A.
 RP TISSUE=LUNG;
 RC MEDLINE-97349118; PubMed-9205122;
 RX Yamada T., Nezu J., Shimane M., Hirata Y.;
 RA "Molecular cloning of a novel vascular endothelial growth factor,
 RT VEGF-D";
 RL Genomics 42:483-488(1997).
 [3]

RP SEQUENCE FROM N.A.
 RX MEDLINE-98118549; PubMed-9435229;
 RA Achen M.G., Jeltsch M., Kukr E., Maekinen T., Vitelli A., Wilks A.F.,
 RA Allitalo K., Stacker S.A.;
 RT "Vascular endothelial growth factor D (VEGF-D) is a ligand for the
 RT tyrosine kinases VEGF receptor 2 (Flk1) and VEGF receptor 3 (Flt4).";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:548-553(1998).
 DR EMBL; Y12864; CAA73371.1; JOINED.
 DR EMBL; Y12865; CAA73371.1; JOINED.
 DR EMBL; Y12866; CAA73371.1; JOINED.
 DR EMBL; Y12867; CAA73371.1; JOINED.
 DR EMBL; Y12868; CAA73371.1; JOINED.
 DR EMBL; Y12869; CAA73371.1; JOINED.
 DR EMBL; Y12870; CAA73371.1; JOINED.
 DR EMBL; D89630; BAA24264.1;
 DR EMBL; AJ000185; CAA03942.1;
 DR EMBL; Y12863; CAA73370.1;
 DR HSP: P15692; IVP.
 DR InterPro: IPR000072;
 DR Pfam: PF00341; PDGF; 1.
 DR ProDom: PD001629; PDGF; 1.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS00278; PDGF_2; 1.
 DR SMART: SM00141; PDGF; 1.
 SQ SEQUENCE 354 AA; 40444 MW; 2048D769D735173E CRC64;

Query Match 20.7%; Score 234; DB 4; Length 354;
 Best Local Similarity 28.8%; Pred. No. 1,7e-17;
 Matches 51; Conservative 24; Mismatches 46; Indels 56; Gaps 6;

OY 8 OCCAANKTCPTNYMNNHICRLAQEDMFSSDAGDSDTDFHIDGPNKELDEETCCV 67
 Db 214 RCHSKRLPDIIMDSNCKVLE-----ENPLATED-----HSHLOE----- 254
 OY 68 CARGLRAPASGPRKELDRNSCQVCNKKLFPSQGANREPDENTCQVCYKTRPNQPLN 127
 Db 255 -----PALGPR-----MMFEDRCVCCKTPCKLIDH 284
 OY 128 PGKAC-ECTESPQCKLLGKKFHOTCSC-----YRRPCTNRKACGEFSEYSEE 177
 Db 285 PKNKSCFECKESLETCQCKHKLFPDTCGCEBRCPRHTPCASGKTAACKHCFPRE 341

RESULT 7
 P97946 PRELIMINARY: PRT: 358 AA.
 ID P97946:
 AC P97946:
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR D (C-FOS INDUCED GROWTH FACTOR).
 GN VEGF-D OR FIGF.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=97030254; PubMed=8876195;
RA Orlandini M., Marconini L., Ferruzzi R., Oliviero S.;
RT "Identification of a c-fos-induced gene that is related to the
RT platelet-derived growth factor/vascular endothelial growth factor
RT family.";
RT Proc. Natl. Acad. Sci. U.S.A. 93:11675-11675(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RX MEDLINE=97349118; PubMed=9205122;
RA Yamada Y., Nezu J., Shimane M., Hirata Y.;
RT "Molecular cloning of a novel vascular endothelial growth factor,
RT VEGF-D.";
RL Genomics 42:483-488(1997).
DR EMBL: X99572; CAA67892.1; -;
DR EMBL: D89628; BAA14002.1; -;
DR HSSP: P15692; 1VPP.
DR MGD: MGI:108037; F19F.
DR InterPro: IPR000072; -;
DR Pfam: PF00341; PDGF_1.
DR ProDom: PD001629; -; 1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS50278; PDGF_2; 1.
DR SMART: SM00141; PDGF; 1.
SQ SEQUENCE 358 AA; 40908 MW; 6636B17FBF07037C CRC64;

Query Match 19.9%; Score 225; DB 11; Length 358;
Best Local Similarity 28.2%; Pred. No. 1.6e-16;

Matches 50; Conservative 21; Mismatches 50; Indels 56; Gaps 6;

QY 8 OCOAANKTCPTNYMNNHICRCLAQEDFMFSSDAGDSTGDFHICGPKKELDEETCCVC 67
DB 219 ECPHKKLCIPVMDMNTCKCYLD-----ENPLPGED-----HSTYQE----- 258
QY 68 CRAGLRPASCGRPHKELDRNSCOCVCKNKLPPSOGANREPDENTCOCVCKRCPRNQPLN 127
DB 259 -----EPTLCGRH-----MFEDEDRCEVCYKACPCGDLIOH 289
QY 128 PGKAC-ECTESPQCKLKGKKFHQTCSC-----YRRPCTNRQKACRPGSYSE 177
DB 290 PENCSCFECKESLSCCKQHKHIFHPDTCSEDRCPHTTTCASRKPACGKHWRFPE 346
SQ SEQUENCE 346

RESULT 8
035251
ID 035251 PRELIMINARY: PRT: 326 AA.
AC 035251:
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, last sequence update)
DT 01-JAN-2001 (TRENBLrel. 15, last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR D.
GN VEGF-D.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE DAWLEY;
RX MEDLINE=97349118; PubMed=9205122;
RA Yamada Y., Nezu J., Shimane M., Hirata Y.;
RT "Molecular cloning of a novel vascular endothelial growth factor,
RT VEGF-D.";
RL Genomics 42:483-488(1997).
DR EMBL: AF014827; AAB6557.1; -;

DR HSSP: P15692; 1VPP.
DR InterPro: IPR000072; -;
DR Pfam: PF00341; PDGF; 1.
DR ProDom: PD001629; -; 1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS50278; PDGF_2; 1.
DR SMART: SM00141; PDGF; 1.
SQ SEQUENCE 326 AA; 37112 MW; 1261AFA373596C00 CRC64;

Query Match 17.6%; Score 199; DB 11; Length 326;
Best Local Similarity 29.7%; Pred. No. 9.6e-14;

Matches 44; Conservative 16; Mismatches 38; Indels 50; Gaps 5;

QY 8 OCOAANKTCPTNYMNNHICRCLAQEDFMFSSDAGDSTGDFHICGPKKELDEETCCVC 67
DB 219 ECPHKKLCIPVMDMNTCKCYLD-----ENPLPGED-----HSTYQE----- 259
QY 68 CRAGLRPASCGRPHKELDRNSCOCVCKNKLPPSOGANREPDENTCOCVCKRCPRNQPLN 127
DB 260 -----PALCGH-----MFEDEDRCEVCYKACPCGDLIOH 289
QY 128 PGKAC-ECTESPQCKLKGKKFHQTC 154
DB 290 PENCSCFECKESLSCCKQHKHIFHPDTC 317
SQ SEQUENCE 317

RESULT 9
074567
ID 074567 PRELIMINARY: PRT: 704 AA.
AC 074567:
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, last annotation update)
DE Q174. PROTEIN.
GN Q174.
OS Trichoderma harzianum.
OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Trichoderma.
OX NCBI_TaxId=5544;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CECT 2413;
RX MEDLINE=98263335; PubMed=9600944;
RA Rey M., Ohno S.A., Pintor-Toro J.A., Jose A., Llobell A., Benitez T.;
RT "Unexpected homology between inducible cell wall protein Q174 of
RT filamentous fungi and B83 salivary protein of the insect Chironomus.";
RT Proc. Natl. Acad. Sci. U.S.A. 95:6212-6216(1998).
DR EMBL: X95671; CAA64974.1; -;
DR InterPro: IPR000561; -;
DR PROSITE: PS01186; EGF_2; 1.
SQ SEQUENCE 704 AA; 77925 MW; 6341ABDDEC365EBC CRC64;

Query Match 15.2%; Score 171.5; DB 3; Length 704;
Best Local Similarity 22.8%; Pred. No. 1.7e-10;

Matches 58; Conservative 26; Mismatches 67; Indels 103; Gaps 15;

QY 16 CPTNYMNNHICRCLAQEDFMFSSDAGDSTGDFH-ICGPKKEL--DEETCCVC----- 67
DB 78 CPSGYTWNGHQC-----VHDCGKDATWQYGCVCNKGVEYVNPDKTSCPPGOY 127
QY 68 -----CRALRPASCGRPHKELDRNSCOCVCK--NKLFPs-----Q 100
DB 128 WNGKNCQ-----VDCGKDATYDQKCKYCKKHGELYNSTKSCPPGOYNGYACVVD 182
QY 101 CGANREPDENTCOCVCKR-----TCPRNQLPGKAC-----ECTE-- 137
DB 183 CGKAHFQDKQKCCVNNGELYNSTKSCACPGQYFNGKCKVCYPGKWNQKQCVEDC 242
QY 138 -----SPQCKLCK--GKFFH--HQTCSC-----YRRPCTNRQ--KACRPG-- 171
DB 243 GREAHFDYQKCKVCNKGVEYVNSAKTSCPDGOYWNGKQCVCPYGOVFNQKQCVPCDG 302

OY 164 ROKACEGFSYSEVCRVPSY 185
 DB 442 --KGCENGFCDAPKSCNDGY 461

RESULT 12

AC 090994 PRELIMINARY; PRT: 1532 AA.
 ID 090994
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
 DE 190 KDA TENASCIN PRECURSOR.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Spring J.;
 RL Submitted (JUL-1989) to the EMBL/Genbank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90030407; PubMed=2478295;
 RA Spring J., Beck K.F., Chiquet-Ehrismann R.;
 RT recombinant tenascin fragments."
 RL Cell 59:325-334(1989).
 DR EMBL; M23121; AAA49085.1; .
 DR HSSP; P24821; ITEN.
 DR InterPro; IPR000561; .
 DR InterPro; IPR001777; .
 DR InterPro; IPR002181; .
 DR Pfam; PF00008; EGF_13.
 DR Pfam; PF00041; fn3; 8.
 DR Pfam; PF00147; fibrinogen_C_1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_14.
 DR PROSITE; PS01186; EGF_2; 14.
 DR SMART; SM00060; FN3; 1.
 KW EGF-like domain; Glycoprotein; Signal.
 FT SIGNAL 1 33
 FT CHAIN 34 1532 POTENTIAL.
 FT CHAIN 34 1224 POTENTIAL.
 FT CHAIN 34 1042 POTENTIAL.
 FT CHAIN 1043 1532 POTENTIAL.
 FT SEQUENCE 1532 AA; 168632 MW; 1E62B9ED8AFC3ACB CRC64;

Query Match 13.1%; Score 148; DB 13; Length 1532;
 Best Local Similarity 24.2%; Pred. No. 1, 1e-07;

Matches 61; Conservative 20; Mismatches 79; Indels 92; Gaps 18;

OY 7 POCQAANKTCPTNYMNNHIC--RCLAQEDF-----MFSSDAGDS--TDG----F 49
 DB 185 PNC--SEPCAPRNCL-NRGLCYRGKICICEGFTGEDCSQAACPSDNDGKCVGCVCF 241
 OY 50 HDICGNKELDEETC-----QCVCRAGLRPASCG----PHKELDRNSC---Q 89
 DB 242 EGYTGP--DCGELCPHGCGIHGRVCVCHGEFTGEDCNELCPNCHNRGRVND 299
 OY 90 CYCKNKLFPSSCG--ANREFDENTC--QCVCRA--TCPRNOPLN----P 128
 DB 300 CVCDEGYTGEDGELICPDNCFDRGRCINGTCFCEGYTGEDCGELTTCNNCNGNRCEN 359
 OY 129 GKCAC-----ECTES--PQCKLLKGGKFHHQJSCY-----RRP 160
 DB 360 GLCVCHGEFVGDDCSQKRCPKDCNNRNGHCVDGR-CVCHGGLYGEDCGELRCPNDCHNRGR 418
 OY 161 CTNRQKACEPGF 172
 DB 419 CINGOCVCDEGF 430

RESULT 13

AC 090995 PRELIMINARY; PRT: 1714 AA.
 ID 090995
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
 DE 200 KDA TENASCIN PRECURSOR.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Spring J.;
 RL Submitted (JUL-1989) to the EMBL/Genbank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90030407; PubMed=2478295;
 RA Spring J., Beck K.F., Chiquet-Ehrismann R.;
 RT recombinant tenascin fragments."
 RL Cell 59:325-334(1989).
 DR EMBL; M23121; AAA49084.1; .
 DR HSSP; P24821; ITEN.
 DR InterPro; IPR000561; .
 DR InterPro; IPR001777; .
 DR InterPro; IPR002181; .
 DR Pfam; PF00008; EGF_13.
 DR Pfam; PF00041; fn3; 10.
 DR Pfam; PF00147; fibrinogen_C_1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_14.
 DR PROSITE; PS01186; EGF_2; 14.
 DR SMART; SM00186; EGF_1.
 KW EGF-like domain; Glycoprotein; Signal.
 FT SIGNAL 1 33
 FT CHAIN 34 1532 POTENTIAL.
 FT CHAIN 34 1224 POTENTIAL.
 FT CHAIN 34 1714 POTENTIAL.
 FT CHAIN 1225 1714 POTENTIAL.
 FT SEQUENCE 1714 AA; 188515 MW; A6FB0CFD623CDE53 CRC64;

Query Match 13.1%; Score 148; DB 13; Length 1714;
 Best Local Similarity 24.2%; Pred. No. 1, 2e-07;

Matches 61; Conservative 20; Mismatches 79; Indels 92; Gaps 18;

OY 7 POCQAANKTCPTNYMNNHIC--RCLAQEDF-----MFSSDAGDS--TDG----F 49
 DB 185 PNC--SEPCAPRNCL-NRGLCYRGKICICEGFTGEDCSQAACPSDNDGKCVGCVCF 241
 OY 50 HDICGNKELDEETC-----QCVCRAGLRPASCG----PHKELDRNSC---Q 89
 DB 242 EGYTGP--DCGELCPHGCGIHGRVCVCHGEFTGEDCNELCPNCHNRGRVND 299
 OY 90 CYCKNKLFPSSCG--ANREFDENTC--QCVCRA--TCPRNOPLN----P 128
 DB 300 CVCDEGYTGEDGELICPDNCFDRGRCINGTCFCEGYTGEDCGELTTCNNCNGNRCEN 359
 OY 129 GKCAC-----ECTES--PQCKLLKGGKFHHQJSCY-----RRP 160
 DB 360 GLCVCHGEFVGDDCSQKRCPKDCNNRNGHCVDGR-CVCHGGLYGEDCGELRCPNDCHNRGR 418
 OY 161 CTNRQKACEPGF 172
 DB 419 CINGOCVCDEGF 430

RESULT 14

Q9VJ05 PRELIMINARY; PRT: 543 AA.

Q9VJ05
AC Q9VJ05
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAR-2001 (TREMBLrel. 13, Last sequence update)
DE BG:DS00180.10 PROTEIN.
GN BG:DS00180.10 OR CG8942.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Chape M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abill J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durkin K.J., Evangelista C.C., Ferraz C., Fertlera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laspo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merulov G., Mishina N.Y., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon S., Nusskern D.R., Paule J.M.,
RA Palazzolo M., Peltan G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Relbert C., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Ventler E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL: AE003642; AAF5364.1; -
DR HSSP: P35555; 1EMN.
DR FLYbase: FBgn0028545; BG:DS00180.10.
DR InterPro: IPR000561; -
DR InterPro: IPR001304; -
DR PROSITE: PS00615; C-TYPE-LECTIN_1; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 7.
DR SMART: SM00181; EGF_1.
KW EGF-like domain; Glycoprotein.
SQ SEQUENCE 543 AA; 58562 MW; F419FDAFC4AD0D2A CRC64;

Query Match 12.8%; Score 145; DB 5; Length 543;
Best Local Similarity 23.4%; Pred. No. 9.6e-08;
Matches 57; Conservative 19; Mismatches 96; Indels 70; Gaps 15;
QY 1 SUPARTPOCOA-----ANKTCPTNYMNNHICR-----CLAQEDPMFSSDAGD-----DS 45

Db 31 SSPNTC-SCNAGYGIDCHPVCPT-VCGKNRCDRPGVSCQNGYKRTSPSDNCLPVCCK 88

QY 46 TDFGFDICPPKNKELDEFTQCVCRCAGLRAPSCGP-----HKELDRSCQCCVCKNKLFPQ 100

Db 89 ECGHHSFC-----SEPKCECEPEYKVGNGTYPDPGKNNKSNKSPIC-----PRD 136

QY 101 CGAN-REFDENTCOC-----VCKRTCPRN-OPLPNGKACE----- 134

Db 137 CGQNSRCVAPGVCCEENGAGDGGTNCNRPVCS-TCPENGCLISPGVCYCKRGYWRNDL 195

QY 135 CTEPQCKLLKKRKHQTCSCY-----RRPCTNR-QKACEPGFSYSEVCRCVPSY 185

Db 196 CQPCHEKCSDMNHCVPANQCECFPGYESSGADKKCVPKCSKCTNCFAPAPFCVCSIGY 255

QY 186 WKRP 189

Db 256 QMGP 259

RESULT 15

Q9NKB8 PRELIMINARY; PRT: 620 AA.

Q9NKB8
AC Q9NKB8
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE BG:DS00180.10 PROTEIN.
GN BG:DS00180.10 OR CG8942.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-Y, CN BW SP;
RX MEDLINE-99403001; PubMed-10471707;
RA Ashburner M., Miera S., Roote J., Lewis S.E., Blazej R., Davis T.,
RA Doyle C., Galie R., George R., Harris N., Hartell G., Harvey D.,
RA Hong L., Houston K., Hoskins R., Johnson G., Martin G., Moshrefi A.,
RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitehead K.,
RA Celniker S., Rubin G.M.;
RT "An exploration of the sequence of a 2.9-Mb region of the genome of Drosophila melanogaster: the Adh region.";
RL Genetics 153:179-219(1999).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-Y, CN BW SP;
RA Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,
RA Butenhoff C., Chape M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA Farfan D.E., Galie R., George R.A., Harris N.L., Hoskins R.A.,
RA Houston K.A., Humastil S.R., Kaira K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lomolan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA Nixon K., Paule J.M., Park S., Pfeiffer B., Poon L., Sequerra A.,
RA Sethi H., Snit E., Svitskas R.R., Wan K.H., Weinburg T., Zhang R.,
RA Zieran L.L., Rubin G.M.;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AE003408; AAF4484.1; -
DR FLYbase: FBgn0028545; BG:DS00180.10.
DR InterPro: IPR000561; -
DR InterPro: IPR001304; -
DR PROSITE: PS00615; C-TYPE-LECTIN_1; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 7.
DR SMART: SM00181; EGF_1.
KW EGF-like domain; Glycoprotein.
SQ SEQUENCE 620 AA; 67953 MW; 29959AC3C8A51D43 CRC64;

Query Match 12.8%; Score 145; DB 5; Length 620;
Best Local Similarity 23.4%; Pred. No. 1.1e-07;
Matches 57; Conservative 19; Mismatches 98; Indels 70; Gaps 15;


```

XX 11-JUN-1998.
PD
XX
XX
PF 20-NOV-1997; 97MO-US20888.
XX
XX 18-SEP-1997; 97US-0933455.
PR 06-DEC-1996; 96US-0759657.
XX
PA (ZYMO ) ZYMOGENETICS INC.
XX
XX Conklin DC, Gilbert T, Hart CE, Nygaard S, Sheppard PO;
XX
XX WPI; 1998-333256/29.
DR N-PSDB; AAV32823.
XX
XX
PT New isolated vascular endothelial growth factor - used to develop
PT products for treating e.g. wounds, burns, myocardial infarction,
PT tumours, psoriasis, arthritis, restenosis or organ transplants
XX
XX
PS Claim 1; Pages 53-54; 77pp; English.
XX
XX The present sequence represents a human zvegf2 growth factor encoded
CC by the zvegf2 cDNA which was isolated from a human heart cDNA library.
CC zvegf2 protein in a dimeric form acts as a mitogen for fibroblasts or
CC smooth muscle cells. zvegf2 is claimed to be useful for stimulating the
CC revascularisation of tissue or the re-endothelialisation of vascular
CC tissue. zvegf2 is particularly claimed to be useful for the treatment
CC of full-thickness skin wounds, including venous stasis ulcers and
CC diabetic ulcers. The zvegf2 protein is also claimed to be useful as an
CC additive in tissue adhesives for promoting revascularisation of the
CC healing tissue. Antagonists against zvegf2 can be used to block its
CC mitogenic, chemotactic and angiogenic effects. The antagonists may
CC therefore be useful for reducing growth of solid tumours by inhibiting
CC neovascularisation of the developing tumour or by directly blocking
CC tumour cell growth. In the treatment of diabetic retinopathy, psoriasis,
CC arthritis, and scleroderma.
XX
XX
SQ Sequence 354 AA:

```

```

Query Match 100.0%; Score 1963; DB 19; Length 354;
Best Local Similarity 100.0%; Pred. No.3.5e-149;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MYREVVVNFMMLYVOLVOGSSNENHGPVKRSSOSTLERSEOOIRAASSLEELLRTITHESE 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 myrewwvvvfmmlylvqlvgssnehgvpvkrsqstlerseqgiraassleellritlse 60

QY 61 DWKIMRCRLKSTSMDSASASRSTRPAATFYDITLTKLVIDEMORTCCSPRETCVEV 120
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 dwkimrcrlkfstmsdsasasrstrfaatfydletlklvdeewgrftqscpretcev 120

QY 121 ASELGKSTNFFKPPCNVNRCCGCCNEESLICMNTSTVSIKOLRETSVPLTSVPELTP 180
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 aselgkstnffkppcnvnrccgcsneeslicmntstviskqlleisvpltsvpeelp 180

QY 181 VKVANHGTGCKLPAPRNPISIRSIQIPEDEDCSSSKKLCIPDMLMDNKKCVQEE 240
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 vkanhgtagcklparnpisirsirqipeedcshskklicpdlmvdnknckcvlqee 240

QY 241 NPLAGTEDSHLOEPALCGPHMFMDEDCRCVCCTPCPKDLIOHPKNCSEFECKESLETC 300
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 nplagtedshlgepalcghmfmdeedrccevcctpcpkdlqhpknscfeckesleetc 300

QY 301 CQHKKLFHPPTSCEDKCFPHTRPCASGKTAACKHCFPEKRAAGCPHRRKP 354
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 cqkhklfhptscedkcfphtrpcasgktaackhcfpekraagcphrrknp 354

```

```

RESULT 2
AAW53241
ID AAW53241 standard; Protein; 354 AA.
XX

```

```

AC AAW53241;
XX
XX 03-AUG-1998 (first entry)
DE Homo sapiens vascular endothelial growth factor D (VEGF-D).
XX
XX vascular endothelial growth factor; VEGF-D; angiogenesis;
XX modification; acceleration; wound healing; tissue; organ;
XX transplants; collateral circulation; infarction; arterial stenosis;
XX coronary artery disease; inhibition; cancer; treatment;
XX diabetic retinopathy; lung disorders; blood circulation;
XX gaseous exchange; chronic obstructive airway disease;
XX intestinal malabsorptive syndrome; biopsy; metastatic risk;
XX detection; diagnosis; congestive heart failure.
XX
XX Homo sapiens.
OS
XX
XX WO9807832-A1.
PN
XX
XX 26-FEB-1998.
PD
XX
XX 21-AUG-1997; 97MO-US14696.
PF
XX
XX 01-JUL-1997; 97US-0051426.
PR
XX 23-AUG-1996; 96AU-0001825.
PR
XX 23-AUG-1996; 96US-0023751.
PR
XX 11-NOV-1996; 96AU-0003554.
PR
XX 14-NOV-1996; 96US-0031097.
PR
XX 05-FEB-1997; 97AU-0004954.
PR
XX 10-FEB-1997; 97US-0038814.
PR
XX 19-JUN-1997; 97AU-0007435.
XX
XX (LUDWIG-) LUDWIG INST CANCER RES.
PA (UYHE-) UNIV HELSINKI LICENSING LTD.
XX
XX Achen MG, Allitalo K, Stackner SA, Wilks AF;
XX
XX WPI; 1998-179057/16.
DR
XX
XX N-PSDB; AAV20807.
DR
XX
XX New isolated vascular endothelial growth factor-D - used to develop
PT products for use in e.g. modifying angiogenesis or treating lung,
PT heart or intestinal disorders
XX
XX
XX Claim 16; Pages 60-61; 101pp; English.
XX
XX The sequence is that of human lung vascular endothelial growth factor
CC D (VEGF-D). VEGF-D can be used for e.g. acceleration of angiogenesis
CC in wound healing, tissue or organ transplantation, or to establish
CC collateral circulation in tissue infarction or arterial stenosis,
CC such as coronary artery disease, and inhibition of angiogenesis in
CC the treatment of cancer or of diabetic retinopathy. It can also be
CC used in the treatment of lung disorders to improve blood circulation
CC in the lung and/or gaseous exchange between the lungs and the blood
CC stream or to improve blood circulation to the heart and O2 gas
CC permeability in cases of cardiac insufficiency, to improve blood
CC flow and gaseous exchange in chronic obstructive airway disease,
CC or to treat malabsorptive syndromes in the intestinal tract.
CC Quantitation of VEGF-D in cancer biopsy specimens may be useful
CC as an indicator of future metastatic risk. Antagonists can be used
CC for treating e.g. conditions such as congestive heart failure,
CC involving accumulations of fluid in the lung resulting from
CC increases in vascular permeability. The products can also be used
CC for detection and diagnosis.
XX
XX
SQ Sequence 354 AA:

```

```

Query Match 100.0%; Score 1963; DB 19; Length 354;
Best Local Similarity 100.0%; Pred. No.3.5e-149;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MYREVVVNFMMLYVOLVOGSSNENHGPVKRSSOSTLERSEOOIRAASSLEELLRTITHESE 60

```

```
|||||
Db 1 myrewvvvvfmmllyvqlyvgssenehgvkrsqstlerseqdiraassleellritthse 60
QY 61 DKMLWCRRLRLKSFTSMDSRSASHRSTRPAAFFYDIETLKVIDEEMQRQCSPRETCVEY 120
  61 dwklwrcrlrlksftsmdsrsashrstriaatfydieltklvideewqrqcspretcvev 120
Db 121 ASFLGKSTNTPFRPCVNVNFRGCGCCNESLJCMNSTSYISKQLEISVPLTSVELVP 180
  121 aselgkstntffkpcvvnfrcgccneesllcmnstsyiskqlteisvpltsvelvp 180
QY 181 VKVANTGCKCLPTAPRHYSIIRRSIQIPEEDRCSHKKLCPIDMLMDSNCKCVLOEE 240
  181 vkvanhtgckclptaprhpysiiirrsiqipeedrcshskklcpidmlwsnckcvlgee 240
Db 241 NPLAGTSDHSHLOEPALCGPHMFMDEDRCQCVKTPCPKDLIOHPKNCSCFECKESLETC 300
  241 nplagtedshshlqepalcgphmmfdeedrccevcoktpcpkdl1qhpknscfeckesle 300
QY 301 CQKHKLFPDTCSCEDRCPEFHTRPCASGKTKACAKHCRFPKERRAAGPHSRKMP 354
  301 cqkhklfhpdtscsedrcpfehtrpcasgktaacakhcrfpkexraagphsrkmp 354
Db
```

RESULT 3

```
AAW44293
ID AAW44293 standard; Protein; 354 AA.
AC AAW44293;
XX
XX
XX 22-JUN-1998 (first entry)
DT
XX
XX Human vascular endothelial growth factor D.
DE
XX
XX Human; vascular endothelial growth factor D; VEGF-D; gene therapy;
KW inflammation; oedema.
XX
XX Homo sapiens.
OS
XX
XX MO9802543-A1.
PN
XX
XX 22-JAN-1998.
PD
XX
XX 15-JUL-1997; 97WO-JP02456.
PF
XX
XX 15-JUL-1996; 96JP-0185216.
PR
XX
XX (CHUGAI) CHUGAI RES INST MOLECULAR MEDICINE INC.
PA
XX
XX Hirata Y, Nezu J;
PI
XX
XX WPI: 1998-110591/10.
DR
XX
XX N-PSDB; AAV15156.
DR
XX
XX VEGF-D protein encoded by DNA - useful for, e.g. gene therapy and
PT treating oedema
PS
XX
XX Claim 1; Page 18-20; 52pp; Japanese.
XX
XX
XX The present sequence represents human vascular endothelial growth factor
CC D (VEGF-D). The VEGF-D protein, compounds and antibodies, which can bind
CC the protein, may be useful in, e.g. gene therapy and in treatment of
CC inflammation and oedema. Vectors, containing the VEGF-D DNA, and VEGF-D
CC DNA sequences may be used for screening for the compounds which bind to
CC the VEGF-D protein.
XX
XX
XX Sequence 354 AA;
SQ
```

Query Match 100.0%; Score 1963; DB 19; Length 354;
Best Local Similarity 100.0%; Pred. No. 3.5e-149;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MYREWVVVVFMMLYVOLYVGSSNEHGVKRRSSOSTLERSEQDIRAASSHEELLRIITHSE 60
  1 myrewvvvvfmmllyvqlyvgssenehgvkrsqstlerseqdiraassleellritthse 60
Db 61 DKMLWCRRLRLKSFTSMDSRSASHRSTRPAAFFYDIETLKVIDEEMQRQCSPRETCVEY 120
  61 dwklwrcrlrlksftsmdsrsashrstriaatfydieltklvideewqrqcspretcvev 120
Db 61 dwklwrcrlrlksftsmdsrsashrstriaatfydieltklvideewqrqcspretcvev 120
QY 121 ASFLGKSTNTPFRPCVNVNFRGCGCCNESLJCMNSTSYISKQLEISVPLTSVELVP 180
  121 aselgkstntffkpcvvnfrcgccneesllcmnstsyiskqlteisvpltsvelvp 180
Db 121 aselgkstntffkpcvvnfrcgccneesllcmnstsyiskqlteisvpltsvelvp 180
QY 181 VKVANTGCKCLPTAPRHYSIIRRSIQIPEEDRCSHKKLCPIDMLMDSNCKCVLOEE 240
  181 vkvanhtgckclptaprhpysiiirrsiqipeedrcshskklcpidmlwsnckcvlgee 240
Db 181 vkvanhtgckclptaprhpysiiirrsiqipeedrcshskklcpidmlwsnckcvlgee 240
QY 241 NPLAGTSDHSHLOEPALCGPHMFMDEDRCQCVKTPCPKDLIOHPKNCSCFECKESLETC 300
  241 nplagtedshshlqepalcgphmmfdeedrccevcoktpcpkdl1qhpknscfeckesle 300
Db 241 nplagtedshshlqepalcgphmmfdeedrccevcoktpcpkdl1qhpknscfeckesle 300
QY 301 CQKHKLFPDTCSCEDRCPEFHTRPCASGKTKACAKHCRFPKERRAAGPHSRKMP 354
  301 cqkhklfhpdtscsedrcpfehtrpcasgktaacakhcrfpkexraagphsrkmp 354
Db
```

RESULT 4

```
AAB10649
ID AAB10649 standard; Protein; 354 AA.
AC AAB10649;
XX
XX
XX 19-JAN-2001 (first entry)
DT
XX
XX Human VEGD protein.
DE
XX
XX VEGF-X; vascular endothelial growth factor; human; vulnery; cytostatic;
KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
KW venous sore; diabetic ulcer; burns; skin graft growth; VEGD.
XX
XX Homo sapiens.
OS
XX
XX WO200037641-A2.
PN
XX
XX 29-JUN-2000.
PD
XX
XX 21-DEC-1999; 99WO-US30503.
PF
XX
XX 22-DEC-1998; 98GB-0028377.
PR
XX
XX 18-MAR-1999; 99US-0124967.
PR
XX
XX 08-NOV-1999; 99US-0164131.
XX
XX (JANC) JANSSEN PHARM NV.
PA
XX
XX Gordon RD, Sprengel JJ, Von JR, Dijkmans JH, Goslowska A;
PI
XX
XX Dhanaraj SN, Xu J;
XX
XX WPI: 2000-442669/38.
XX
XX New vascular endothelial growth factor protein, useful for treating or
CC preventing diseases associated with inappropriate angiogenesis activity
CC such as cancer, rheumatoid arthritis, psoriasis and wounds -
XX
XX Disclosure; Fig 11; 127pp; English.
XX
XX This invention describes a novel vascular endothelial growth factor-X
CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
CC vulnery, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
CC antidiabetic activity and acts as an angiogenesis and vascularization
CC regulator. An antisense molecule of the invention is useful for treating
CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
```

CC retinopathy by inhibiting angiogenic activity or inappropriate
 CC vascularization including formation and proliferation of new blood
 CC vessels, growth and development of tissues, tissue regeneration and organ
 CC and tissue repair in a subject. The products of the invention are useful
 CC for preparing medicaments for treating wounds such as dermal ulcers,
 CC pressure sores, venous sores, diabetic ulcers and burns and to promote
 CC skin graft growth, tissue repair, proliferation of new blood vessels,
 CC tissue regeneration and organ repair by promoting angiogenic activity or
 CC vascularization. This sequence represents the human VEGD protein used
 CC to illustrate the method of the invention.
 CC
 XX
 SQ Sequence 354 AA:
 Query Match 100.0%; Score 1963; DB 21; Length 354;
 Best Local Similarity 100.0%; Pred. No. 3.5e-149;
 Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MYREVVVVVFMMLYVQIVOGSSNEHGPVKRSSQSTLERSEQQIRASSLEELLRTTSE 60
 DB 1 myrewvvvvfmmllyvqivgssnehgpkrrsqstlerseqqiraassleellrttshse 60
 QY 61 DWKLMRCRLKLSFTSMDSRSASHRSTRFAATFYDIETLKVIDEEMORTQCSPRETCVEV 120
 DB 61 dwklwrcrlklsftsmdsrsashrstrfaattydieltkvideeqgrtcspretcvev 120
 QY 121 ASELGKSTNFFKPCPVNVNFRCGGCNEESLICMNTSTSYISKQLEISVPLTSVPELVP 180
 DB 121 aselgkstntffkpcpvnvfrcggcneesliomntstsyiskqlfeisvpltsvpeylvp 180
 QY 181 VKVANTGCKCLPTARPHRYSTIIRSIQIPEDRCHSKKLCPIIDMLMNSNCKCYLOEE 240
 DB 181 vkvanthgckclptarphrystiirsiqipedrshskklcpiidmlwnskckvlygee 240
 QY 241 NPLAGTEDSHLOEPALCGPHMFMDEDRCVCYKTPCPKDLIOHPKNSCFECKESLETC 300
 DB 241 nplagtedshlqepalcgphmfmfededrcvcyktpcpkdlighpkncsfekesleetc 300
 QY 301 CQKHKLFPDTCSCEDRCPEHTRPCASGKTAACAKHCRFPKERRAAGPHSRKNP 354
 DB 301 cqkhklfhpdtscedrcpfhtrpcasgktaacakhcrfprekraagphsrknp 354
 RESULT 5
 AAB29049
 ID AAB29049 standard; Protein; 354 AA.
 XX
 AC AAB29049;
 XX
 DT 31-JAN-2001 (first entry)
 XX
 DE Human VEGF-D protein sequence.
 XX
 KW Human; Flt4; fms-like tyrosine kinase 4; lymphoedema;
 KW vascular endothelial growth factor receptor 3; VEGFR-3;
 KW Milroy-Nonne syndrome; lymphoedema praecox; VEGF-D;
 KW vascular endothelial growth factor D.
 XX
 OS Homo sapiens.
 XX
 PN WO200058511-A1.
 XX
 PD 05-OCT-2000.
 XX
 PF 26-MAR-1999; 99WO-US06133.
 XX
 PR 26-MAR-1999; 99WO-US06133.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 PA (UYHE-) UNIV HELSINKI LICENSING LTD OY.
 PA (UYPI-) UNIV PITTSBURGH.
 XX
 PI Ferrell RE, Altalo K, Finegold DN, Karkkainen M;

XX
 DR WPI; 2000-679298/66.
 DR N-PSDB; AAC62407.
 XX
 PT Screening a human subject for increased risk of developing a lymphatic
 PT disorder, comprises assaying a nucleic acid to determine a mutation
 PT altering the sequence of a vascular endothelial growth factor
 PT receptor-3 -
 XX
 PS Disclosure; Page 64-65; 76pp; English.
 XX
 CC The present sequence is the protein sequence of the human vascular
 CC endothelial growth factor D (VEGF-D). It was used to demonstrate the
 CC methods of the invention, which involve the screening of individuals to
 CC determine which vascular endothelial growth factor receptor 3 (VEGFR-3,
 CC also known as Flt4 or fms-like tyrosine kinase 4) alleles they possess
 CC and thus their likelihood of developing hereditary lymphoedema.
 CC Conditions associated with lymphoedema include Milroy-Nonne syndrome,
 CC which is early onset lymphoedema and lymphoedema praecox, which is late
 CC onset.
 CC
 XX
 SQ Sequence 354 AA:
 Query Match 100.0%; Score 1963; DB 21; Length 354;
 Best Local Similarity 100.0%; Pred. No. 3.5e-149;
 Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MYREVVVVVFMMLYVQIVOGSSNEHGPVKRSSQSTLERSEQQIRASSLEELLRTTSE 60
 DB 1 myrewvvvvfmmllyvqivgssnehgpkrrsqstlerseqqiraassleellrttshse 60
 QY 61 DWKLMRCRLKLSFTSMDSRSASHRSTRFAATFYDIETLKVIDEEMORTQCSPRETCVEV 120
 DB 61 dwklwrcrlklsftsmdsrsashrstrfaattydieltkvideeqgrtcspretcvev 120
 QY 121 ASELGKSTNFFKPCPVNVNFRCGGCNEESLICMNTSTSYISKQLEISVPLTSVPELVP 180
 DB 121 aselgkstntffkpcpvnvfrcggcneesliomntstsyiskqlfeisvpltsvpeylvp 180
 QY 181 VKVANTGCKCLPTARPHRYSTIIRSIQIPEDRCHSKKLCPIIDMLMNSNCKCYLOEE 240
 DB 181 vkvanthgckclptarphrystiirsiqipedrshskklcpiidmlwnskckvlygee 240
 QY 241 NPLAGTEDSHLOEPALCGPHMFMDEDRCVCYKTPCPKDLIOHPKNSCFECKESLETC 300
 DB 241 nplagtedshlqepalcgphmfmfededrcvcyktpcpkdlighpkncsfekesleetc 300
 QY 301 CQKHKLFPDTCSCEDRCPEHTRPCASGKTAACAKHCRFPKERRAAGPHSRKNP 354
 DB 301 cqkhklfhpdtscedrcpfhtrpcasgktaacakhcrfprekraagphsrknp 354
 RESULT 6
 AAY70750
 ID AAY70750 standard; Protein; 354 AA.
 XX
 AC AAY70750;
 XX
 DT 17-AUG-2000 (first entry)
 XX
 DE Human prepro-vascular endothelial growth factor D.
 XX
 KW Human; receptor tyrosine kinase; RTK; Flt4; fms-like tyrosine kinase 4;
 KW VEGFR-3; vascular endothelial growth factor receptor-3; chromosome 5q35;
 KW cytostatic; tumour imaging; anti-tumour therapy; treatment; diagnosis;
 KW neoplastic disease; lymphoma; carcinoma; breast; squamous cell; melanoma;
 KW sarcoma; malignancy; VEGF-D; vascular endothelial growth factor D.
 XX
 OS Homo sapiens.
 XX
 PN WO200021560-A1.
 XX

[illegible]

Query Match	100.0%	Score 1963	DB 21	Length 354
Best Local Similarity	100.0%	Pred. No. 3.5e-149		
Matches 354: Conservative	0	Mismatches 0	Indels 0	Gaps 0

PR 16-AUG-1999; 99US-0375248.
 XX (UVEI-) UNIV PITTSBURGH.
 PA (UVEH-) UNIV HELSINKI LITENSING LTD OY.
 PA (LUDM-) LODWIG INST CANCER RES.
 XX
 XX
 PI Altalo K, Ferrell RE, Finegold DN, Karkkainen M;
 DR WPI: 2001-007762/02.
 DR N-PSDB; AAC68954.
 PT Screening a human for an increased risk of developing lymphatic
 PT disorder comprises assaying nucleic acid for alterations in the
 PT sequences expressing vascular endothelial growth factor receptor-3
 XX
 XX Disclosure: Pages 66-67; 99pp; English.
 PS
 XX The present invention relates to a method for screening a human subject
 CC for an increased risk of developing a lymphatic disorder e.g. hereditary
 CC lymphoedema. The method comprises assaying nucleic acid of a human
 CC subject to determine a presence or an absence of a mutation altering the
 CC sequence or expression of vascular endothelial growth factor receptor-3
 CC (VEGFR-3)/fms-like tyrosine kinase 4 (Flt4) allele (see AAC68952 and
 CC AAB37604) and determining an increased risk of developing lymphatic
 CC disorder from presence or absence of the mutation. The presence of a
 CC mutation altering the encoded amino acid sequence or expression of at
 CC least 1 VEGFR-3 allele in the nucleic acid correlates with an increased
 CC risk of developing a lymphatic disorder. Treatment for hereditary
 CC lymphoedema can be provided through the administration of vascular
 CC endothelial growth factor C (VEGF-C) and vascular endothelial growth
 CC factor D VEGF-D genes (via gene therapy) and proteins. The present
 CC sequence is the protein sequence for VEGF-D.
 CC
 XX
 SQ Sequence 354 AA:

Query Match 100.0%; Score 1963; DB 22; Length 354;
 Best Local Similarity 100.0%; Pred. No. 3.5e-149;
 Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYREVVVVVFMMLYQVLYOGSSNEHGPVKRSSOSTLESEDOQIRAASSLEELLRTTSEE 60
 DB 1 myrewvvvvfmmllyvqlyvgssnehgpykrssgstlerseggiraassleellrttsee 60
 QY 61 DKLMRCRLRLKFTSMDSRSASHRSTRFAATFYDIETLKVIDEEMQROGSPRETCVY 120
 DB 61 dklmrcrlrlkftsmdsrsashrstrfaatfydieltkvideeqgrtcsprectcvev 120
 QY 121 ASELGKSTNTFFKPPCVNVFRCGGCCNEESLJCMNTSTSYISKOLFELISVPLTSVPELVP 180
 DB 121 aselgkstntffkppcvnvfrcgccneesljcmntstsyiskqlfelfsvpltsvpelvp 180
 QY 181 VAVANHTGCKCLPTAPRAPHYSITIRRSIQIPEDRCSHKKLCPIDMLWDSNKKCVLOEE 240
 DB 181 vvanhtgckclptapraphysitirrsiqipeedrcshskklcpidmlwdsnkkcvalgee 240
 QY 241 NPLAGEDSHLOEPALCGPHMMFDEDRCECVCKTPCPKDLOHPKNCSEFCCKESLFTC 300
 DB 241 nplagedshlqepalcgphmmfdeedrcvccktpcpkdliohpnkncsefckesleetc 300
 QY 301 CQKHLFHPDTCSCEDRCPFHTRPCASGKTCACAKHCRPFKERRAAGPSRRKNP 354
 DB 301 cqkhlfhpdtscedrcpfhtRPCASGKTCACAKHCRPFKERRAAGPSRRKNP 354

RESULT 10
 AAB70685
 ID AAB70685 standard; protein: 354 AA.
 XX
 AC AAB70685;
 XX
 DT 16-MAY-2001 (first entry)
 XX

DE Human vascular endothelial growth factor D (VEGF-D) protein.
 XX
 KW Human; vascular endothelial growth factor D: VEGF-D; angiogenic;
 KW angiogenesis; c-fos induced growth factor; Fgfr; cardiant; vasotropic;
 KW ischaemic condition; ischaemia; vascular insufficiency;
 KW peripheral vascular disease; coronary artery disease;
 KW myocardial infarction.
 XX
 XX Homo sapiens.
 PN WO200112669-A1.
 XX
 XX 22-FEB-2001.
 PD
 XX
 PF 16-AUG-2000; 2000WO-IB01244.
 PR
 XX 16-AUG-1999; 99US-0149300.
 PA
 XX (UVEI-) UNIV SIENA.
 PI
 XX Oliviero S;
 DR WPI: 2001-202857/20.
 XX
 XX Use of recombinant vascular endothelial growth factor or its
 PT angiogenically active fragment or mutant, for inducing angiogenesis in
 PT vivo or in vitro and for treating myocardial infarction, coronary
 PT artery disease -
 PS
 PS Claim 2; Page 51-52; 55pp; English.

The present invention describes a method for inducing angiogenesis in a
 CC tissue, or area, in need of angiogenesis, in a mammal. The method
 CC comprises administering recombinant vascular endothelial growth factor D
 CC (VEGF-D) or its angiogenically active fragment or mutant (I). The
 CC present sequence represents the human VEGF-D protein, which can be used
 CC in the method of the invention. (I) has cardiant and vasotropic
 CC activities, and is an angiogenesis inducer. The method can be used for
 CC inducing angiogenesis in a tissue in need of angiogenesis, or in an area
 CC in need of angiogenesis in a mammal. (I) is useful for treating various
 CC ischaemic conditions manifested by vascular insufficiency such as
 CC peripheral vascular disease, coronary artery disease or myocardial
 CC infarction.
 CC
 XX
 SQ Sequence 354 AA:

Query Match 98.1%; Score 1926; DB 22; Length 354;
 Best Local Similarity 98.3%; Pred. No. 3.2e-146;
 Matches 348; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MYREVVVVVFMMLYQVLYOGSSNEHGPVKRSSOSTLESEDOQIRAASSLEELLRTTSEE 60
 DB 1 myrewvvvvfmmllyvqlyvgssnehgpykrssgstlerseggiraassleellrttsee 60
 QY 61 DKLMRCRLRLKFTSMDSRSASHRSTRFAATFYDIETLKVIDEEMQROGSPRETCVY 120
 DB 61 dklmrcrlrlkftsmdsrsashrstrfaatfydieltkvideeqgrtcsprectcvev 120
 QY 121 ASELGKSTNTFFKPPCVNVFRCGGCCNEESLJCMNTSTSYISKOLFELISVPLTSVPELVP 180
 DB 121 aselgkstntffkppcvnvfrcgccneesljcmntstsyiskqlfelfsvpltsvpelvp 180
 QY 181 VAVANHTGCKCLPTAPRAPHYSITIRRSIQIPEDRCSHKKLCPIDMLWDSNKKCVLOEE 240
 DB 181 vvanhtgckclptapraphysitirrsiqipeedrcshskklcpidmlwdsnkkcvalgee 240
 QY 241 NPLAGEDSHLOEPALCGPHMMFDEDRCECVCKTPCPKDLOHPKNCSEFCCKESLFTC 300
 DB 241 nplagedshlqepalcgphmmfdeedrcvccktpcpkdliohpnkncsefckesleetc 300
 QY 301 CQKHLFHPDTCSCEDRCPFHTRPCASGKTCACAKHCRPFKERRAAGPSRRKNP 354
 DB 301 cqkhlfhpdtscedrcpfhtRPCASGKTCACAKHCRPFKERRAAGPSRRKNP 354

Db 301 cqhklfhpdcscdrcpfhtpcasgktaacahcfrfpekreagpshrknp 354

RESULT 11

AAW14994
ID AAW14994 standard; Protein; 620 AA.

AAW14994;
XX
XX 05-JUL-1997 (first entry)

XX Human c-Fos induced growth factor (clone HF175 ORF2 product).

XX c-Fos induced growth factor; FIGF; Fos regulated gene;

KW proto-oncogene; lung disorder; cancer; tumour; therapy;

KW antibody; transgenic animal.

XX Homo sapiens.

OS
XX
XX Key Location/Qualifiers

FT Misc-difference 16 /note= "residue 16 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"

FT Misc-difference 26 /note= "residue 26 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"

FT Misc-difference 29 /note= "residue 29 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"

FT Misc-difference 47 /note= "residue 47 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"

FT Misc-difference 71 /note= "residue 71 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"

FT Misc-difference 72 /note= "residue 72 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"

FT Misc-difference 76 /note= "residue translated from ORF2 of HF175
FT is Ile"

FT Misc-difference 136 /note= "residue translated from ORF2 of HF175
FT is Ile"

FT Misc-difference 220 /note= "residue translated from ORF2 of HF175
FT is Ile"

FT Misc-difference 341 /note= "residue translated from ORF2 of HF175
FT is Phe"

FT Misc-difference 344 /note= "residue translated from ORF2 of HF175
FT is Phe"

FT Misc-difference 377 /note= "residue translated from ORF2 of HF175
FT is Leu"

FT Misc-difference 435 /note= "residue 435 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"

FT Misc-difference 486 /note= "residue 486 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"

FT Misc-difference 497 /note= "residue 497 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"

FT Misc-difference 518 /note= "residue 518 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"

FT Misc-difference 541 /note= "residue 541 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"

FT Misc-difference 553 /note= "residue 553 corresponds to an in-frame

FT Misc-difference 557 /note= "stop codon in reading frame 2 of HF175"

FT Misc-difference 562 /note= "residue 562 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"

FT Misc-difference 579 /note= "residue 579 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"

FT Misc-difference 592 /note= "residue 592 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"

FT Misc-difference 593 /note= "residue 593 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"

FT Misc-difference 597 /note= "residue 597 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"

FT Misc-difference 605 /note= "residue 605 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"

FT Misc-difference 608 /note= "residue 608 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"

PN W09712972-A2.

PD 10-APR-1997.

PF 30-SEP-1996; 96WO-1B01113.

PR 13-JUN-1996; 96GB-0012368.

PR 29-SEP-1995; 95GB-0019928.

PA (UYSI-) UNIV SIENA.

PI O1viero S;

XX WPI. 1997-226216/20.

DR N-PSDB; AAT62961.

XX Nucleotide molecule encoding c-Fos induced growth factor protein -
PT useful in therapy, in manufacture of compositions for treatment of
PT developmental disorders and in generation of transgenic animal

PS Claim 3; Fig 2; 64pp; English.

XX 3 Polypeptide sequences (AAW14993-95) are the respective translated
CC sequences of reading frames 1, 2 and 3 of clone HF175 (AAT62961), the
CC human homologue of murine clone F0401 (AAT62960), which codes for a
CC novel c-Fos induced growth factor (FIGF) (see also AAW14992).

CC Examination of the 3 polypeptides indicates that reading frame 2
CC has the longest region free of nonsense codons. FIGF is a c-Fos-
CC dependent autocrine growth factor able to induce cell division
CC entry and, when over-expressed, a transformed phenotype in
CC fibroblasts. It could be implicated in tumours and development.
CC Recombinant FIGF can be produced in transformed host (e.g. CHO
CC cells. It can be used to identify its receptors and in an assay
CC for the identification of agonists and antagonists. Antibodies
CC raised against FIGF can be used to block the function of the
CC protein and thereby inhibit or suppress tumour growth. Transgenic
CC animals expressing FIGF can be generated for use e.g. as models for
CC research.

XX Sequence 620 AA:

SQ

Query Match 97.7%; Score 1917; DB 18; Length 620;
Best Local Similarity 96.0%; Pred. No. 3.2e-145;
Matches 347; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Oy 1 MYREVVVVVFMMLYVOLVOGSSNEHGVPYKRSSOSTLERSEQOIRAASSLLELLRITPSE 60

KW		rheumatoid arthritis; autoimmune disease; allergy; cancer; therapy;
KM		infectious disease; neurodegeneration,
KX		vascular endothelial growth factor-D; VEGF-D.
XX		
OS	Homo sapiens.	
XX		
PN	WO20075163-A1.	
PD	14-DEC-2000.	
XX		
PF	01-JUN-2000; 2000WO-US14925.	
XX		
PR	03-JUN-1999; 99US-0137796.	
XX		
PA	(HOMA-) HUMAN GENOME SCI INC.	
PI	Rosen CA, Ruben SM, Hu J, Cao L;	
XX		
DR	WPt: 2001-071057/08.	
XX	N-PSSDB: AAA91006.	
PT	New nucleic acid encoding angiogenic proteins, useful e.g. for	
PT	promoting healing of wounds and treating peripheral arterial disease,	
PT	critical limb ischaemia or coronary disease -	
XX		
PS	Claim 11; Page 226-227; 244pp; English.	
CC	This sequence is vascular endothelial growth factor-D (VEGF-D),	
CC	which is an angiogenic protein of the invention. The angiogenic proteins	
CC	and the DNA sequences encoding them, are used to prevent, treat or	
CC	ameliorate disease and to detect diseases, or susceptibility, by	
CC	detecting mutations or the presence or amount of angiogenic protein	
CC	expression. Particularly they are used to stimulate wound healing,	
CC	growth of damaged bone and tissue, and for repair of vascular tissue,	
CC	especially peripheral arterial disease, critical limb ischaemia or	
CC	coronary disease. Antagonists of the sequences are used to inhibit	
CC	angiogenesis in tumours and to treat inflammation (where associated with	
CC	increased vascular permeability), diabetic retinopathy, rheumatoid	
CC	arthritis or psoriasis. Agonists are also useful for stimulating	
CC	(lymph)angiogenesis. The proteins are also used to identify specific	
CC	binding agents (potential therapeutic agents) and to raise antibodies.	
CC	The antibodies are useful as therapeutic (ant)agonists; for detection,	
CC	purification and targeting of proteins for in vivo or in vitro diagnosis	
CC	(including imaging) or for therapy (including when linked to e.g. a label	
CC	or cytotoxin); and for immunotyping of cells, e.g. for detecting minimal	
CC	residual disease or haematopoietic progenitor/stem cells. It is also	
CC	contemplated that the sequences might be useful for treating a very wide	
CC	range of other disorders, e.g. autoimmune diseases; allergy; cancer;	
CC	infectious diseases (viral, bacterial, fungal or parasitic);	
CC	neurodegeneration, also as chemotactic agents or for stimulating	
CC	regeneration of the nervous system etc.	
XX		
SQ	Sequence 325 AA;	
	Query Match 91.9%; Score 1804; DB 22; Length 325;	
	Best Local Similarity 100.0%; Pred. No. 1,6e-136;	
	Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	31 RSSOSTLRSSEQIRAAASLEELLITTHSEDMKWRCKRLKLSFTMSMDRSASHSTRRA 90	
Dd		
	2 rssgstlrsseqiraaasleellitltlsedekllwrcrlrlksfmsdrasashstrla 61	
OY	ATFYIETLKIVDEEMORTCSPRETCVAVSELGKSTTFPKPPCVANVFRCGGCENES 150	
Dd		
	62 atfyielkivideewrtqscpretcveaseisgkstctffkpcvwvfrcggcnnes 121	
OY	LICMNTSTSYSKOLFETISVPLTSVPBLPVVANHTGCKCLPTAPRHRYSIIIRSIQT 210	
Dd		
	122 llicmntstsyiskqlfetisypltsvpelpyrvvanhtgckclptaprhysiiirsiglp 181	
OY	EEDRCSHSKKLCIPIDMLDSNKCKCVLOEENPLAGTEDHSHLOEPALCGPHMFDEDRCE 270	
Dd		

Db	182	eeutreshskllopiamlwdsnkcvcqlgeampjagtedshlqepalaJogpnmfiedtce	241
Qy	271	CVCkTPCPKDLIQHFRNCSCEFCESLETCCQKHKLFHPDTCSCEDRCPFHTRPCASGKT	330
Db	242	cvcktcppepdlqphkncscfecslelctccqkhkLfhpdtcscedrpfhtrpcasgkt	301
Qy	331	ACAKHCRPFKEKRAAGPHSRKNP	354
Db	302	acakhcrlpkekraagphsrknp	325
RESULT 14			
ID	AAW53242	standard; Protein: 358	AA.
XX	AAW53242;		
AC	AAW53242;		
XX			
DT	03-AUG-1998	(first entry)	
DE	Mus musculus vascular endothelial growth factor D1 (VEGF-D1).		
XX			
KW	vascular endothelial growth factor; VEGF-D; angiogenesis;		
KW	modification; acceleration; wound healing; tissue; organ;		
KW	transplants; collateral circulation; infarction; arterial stenosis;		
KM	coronary artery disease; inhibition; cancer; treatment;		
KM	diabetic retinopathy; lung disorders; blood circulation;		
KW	gaseous exchange; chronic obstructive airway disease;		
KM	intestinal malabsorptive syndrome; biopsy; metastatic risk;		
KM	detection; diagnosis; congestive heart failure.		
XX			
OS	Mus musculus.		
PN	WO9607832-A1.		
XX			
PD	26-FEB-1998.		
XX			
PF	21-AUG-1997;	97WO-US14696.	
XX			
PR	01-JUL-1997;	97US-0051426.	
PR	23-AUG-1996;	96AU-0001825.	
PR	23-AUG-1996;	96US-0023751.	
PR	11-NOV-1996;	96AU-0003554.	
PR	14-NOV-1996;	96US-0031097.	
PR	05-FEB-1997;	97AU-0004954.	
PR	10-FEB-1997;	97US-0038814.	
PR	19-JUN-1997;	97AU-0007435.	
XX			
PA	(LUDWIG-INST CANCER RES.		
PA	(UYHE-) UNIV HELSINKI LICENSING LTD.		
XX			
PI	Achen MG, Altalo K, Stacker SA, Wilks AF;		
XX			
DR	WPI: 1998-179057/16.		
DR	N-PSDB: AAV20808.		
XX			
PT	New isolated vascular endothelial growth factor-D - used to develop		
PT	products for use in e.g. modifying angiogenesis or treating lung,		
PT	heart or intestinal disorders		
XX			
PS	Claim 16; Pages 63-64; 101pp; English.		
XX			
CC	The sequence is that of mouse lung vascular endothelial growth factor		
CC	D1 (VEGF-D1). VEGF-D1 can be used for e.g. acceleration of angiogenesis		
CC	in wound healing, tissue or organ transplantation, or to establish		
CC	collateral circulation in tissue infarction or arterial stenosis,		
CC	such as coronary artery disease, and inhibition of angiogenesis in		
CC	the treatment of cancer or of diabetic retinopathy. It can also be		
CC	used in the treatment of lung disorders to improve blood circulation		
CC	in the lung and/or gaseous exchange between the lungs and the blood		
CC	stream or to improve blood circulation to the heart and O2 gas		
CC	permeability in cases of cardiac insufficiency, to improve blood		
CC	flow and gaseous exchange in chronic obstructive airway disease.		
CC	or to treat malabsorptive syndromes in the intestinal tract.		

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using .sw model

Run on: October 17, 2001, 14:47:11 ; Search time 62.93 seconds

(without alignments)
115.827 Million cell updates/sec

Title: US-09-427-657-4

Perfect score: 1963
Sequence: 1 MYREVVVVVFMMLYQLVQ.....HCRPKKRAAGPHSRKNP 354

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents-AA:*

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTCUTS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Dackfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1963	100.0	354	US-08-915-795-5	Sequence 5, Appl1
2	1804	91.9	325	US-08-915-795-3	Sequence 3, Appl1
3	1675	85.3	358	US-08-915-795-8	Sequence 8, Appl1
4	1522	77.5	321	US-08-915-795-9	Sequence 9, Appl1
5	704.5	35.9	419	US-08-999-811-2	Sequence 2, Appl1
6	704.5	35.9	419	US-09-042-105-2	Sequence 2, Appl1
7	704.5	35.9	419	US-09-042-105-18	Sequence 18, Appl1
8	704.5	35.9	419	US-08-795-430-8	Sequence 8, Appl1
9	704.5	35.9	419	US-08-510-133A-35	Sequence 35, Appl1
10	704.5	35.9	419	PCT-US96-09001-2	Sequence 11, Appl1
11	696	35.5	415	US-08-795-430-11	Sequence 13, Appl1
12	677	34.5	418	US-08-795-430-13	Sequence 4, Appl1
13	664.5	33.9	350	US-08-999-811-4	Sequence 2, Appl1
14	664.5	33.9	350	US-08-824-996-2	Sequence 4, Appl1
15	664.5	33.9	350	US-09-042-105-4	Sequence 33, Appl1
16	664.5	33.9	350	US-08-510-133A-33	Sequence 33, Appl1
17	664.5	33.9	350	US-08-585-895-33	Sequence 33, Appl1
18	204	10.4	165	US-08-585-895-18	Sequence 33, Appl1
19	204	10.4	165	5219739-19	Sequence 33, Appl1
20	204	10.4	191	US-08-567-200A-2	Sequence 2, Appl1
21	204	10.4	191	US-08-807-992B-2	Sequence 2, Appl1
22	204	10.4	191	US-08-691-794-2	Sequence 2, Appl1
23	204	10.4	191	US-08-795-430-56	Sequence 56, Appl1
24	204	10.4	191	5332671-4	Sequence 56, Appl1
25	200.5	10.2	231	PCT-US96-09001-10	Sequence 10, Appl1
26	200.5	10.2	232	US-08-999-811-7	Sequence 7, Appl1
27	200.5	10.2	232	US-08-807-992B-4	Sequence 4, Appl1

28	200.5	10.2	232	US-09-042-105-7	Sequence 7, Appl1
29	200	10.2	164	5194596-17	Patent No. 5194596
30	200	10.2	164	5219739-17	Patent No. 5219739
31	200	10.2	164	5219739-18	Patent No. 5219739
32	199.5	10.2	190	5332671-3	Patent No. 5332671
33	199.5	10.2	214	5240848-11	Patent No. 5240848
34	197.5	10.1	215	5219739-22	Patent No. 5219739
35	197.5	10.1	232	US-08-824-996-9	Sequence 9, Appl1
36	196.5	10.0	189	US-08-469-427A-15	Sequence 15, Appl1
37	196	10.0	215	US-08-586-039B-49	Sequence 49, Appl1
38	196	10.0	215	5240848-7	Patent No. 5240848
39	195	9.9	215	US-08-807-992B-3	Sequence 3, Appl1
40	194	9.9	190	US-08-586-039B-31	Sequence 31, Appl1
41	194	9.9	214	US-08-586-039B-35	Sequence 35, Appl1
42	189	9.6	190	US-08-569-063C-20	Sequence 20, Appl1
43	176.5	9.0	188	US-08-469-427A-5	Sequence 5, Appl1
44	176.5	9.0	188	US-08-609-443B-5	Sequence 5, Appl1
45	176.5	9.0	188	US-08-569-063C-5	Sequence 5, Appl1

ALIGNMENTS

RESULT 1
US-08-915-795-5
; Sequence 5, Application US/08915795
; Patent No. 6235713
; GENERAL INFORMATION:
; APPLICANT: Marc G. ACHEN
; APPLICANT: Andrew F. WILKS
; APPLICANT: Steven A. STACKER
; APPLICANT: Karl ALTRALDO
; TITLE OF INVENTION: GROWTH FACTOR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
; STREET: 1200 G Street, NW, Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: United States of America
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,795
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26, 269
; REFERENCE/DOCKET NUMBER: 1064/42983
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 354 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: Human Lung
; US-08-915-795-5
Query Match 100.0%; Score 1963; DB 4; Length 354;
Best Local Similarity 100.0%; Pred. No. 7.9e-172;

	Matches	354	Conservative	0	Mismatches	0	Indels	0	Gaps	0
Qy	1	MYREVVVVVAVFMALYYVLQVQSSNEHGPKRRSSQSTLERSQOIRAASSLIELRLTHSE	60							
Db	1	MYREVVVVVAVFMALYYVLQVQSSNEHGPKRRSSQSTLERSQOIRAASSLIELRLTHSE	60							
Qy	61	DKMLRCRLRLCSPTSMDSRSASHSTRPAATFYDIETLKYVIDEEMQRQCSPRETCVAV	120							
Db	61	DKMLRCRLRLCSPTSMDSRSASHSTRPAATFYDIETLKYVIDEEMQRQCSPRETCVAV	120							
Qy	121	ASELCKSTNTFFKPPCVANVFRCGGCGNEBSLLCMNTSTYSISKOLFELISVPLTSVELVP	180							
Db	121	ASELCKSTNTFFKPPCVANVFRCGGCGNEBSLLCMNTSTYSISKOLFELISVPLTSVELVP	180							
Qy	181	VKVVANHTGCKCLPTAPRHPYSIIIRSIQIPEEDRCSHKKLCPIDMLMDSNCKCVLOEE	240							
Db	181	VKVVANHTGCKCLPTAPRHPYSIIIRSIQIPEEDRCSHKKLCPIDMLMDSNCKCVLOEE	240							
Qy	241	NPLAETEDHSHLOEPALCGPHMFPEDRCEYCKITPCPRDLIQHPRKNSCFECKSLETC	300							
Db	241	NPLAETEDHSHLOEPALCGPHMFPEDRCEYCKITPCPRDLIQHPRKNSCFECKSLETC	300							
Qy	301	COKHLTFHPDDTSCEDRCDFHTRPCASGKTAAKRCRPFKEKRAAGPSRRNP	354							
Db	301	COKHLTFHPDDTSCEDRCDFHTRPCASGKTAAKRCRPFKEKRAAGPSRRNP	354							

Query Match	91.9%	Score 1804	DB 4	Length 325
Best Local Similarity	100.0%	Pred. No. 2,4e-157		
Matches	324	Conservative	0	Mismatches 0; Indels 0; Gaps 0
QY	31	RSOSLTLESEOOIRASSLEBELLRITTHSEDMKLMRCRLRLKTSFMSDRSASHNSTRA	90	
Db	2	RSOSLTLESEOOIRASSLEBELLRITTHSEDMKLMRCRLRLKTSFMSDRSASHNSTRA	61	
QY	91	AFVFYDIETLKVIDEEMQRFQSPRETCVAVASELGKSTPTFFKPCPVANFRCGGCCNEES	150	
Db	62	AFVFYDIETLKVIDEEMQRFQSPRETCVAVASELGKSTPTFFKPCPVANFRCGGCCNEES	121	
QY	151	LICAMTSTSYISKOLPEISVPLTSPELVPVAVANHTGCKCPLTAPRHYSTIIRNSTQIP	210	
Db	122	LICAMTSTSYISKOLPEISVPLTSPELVPVAVANHTGCKCPLTAPRHYSTIIRNSTQIP	181	
QY	211	EEDRCSHSHKRLCPIDMLMDSNKKCVLQENPLAGTGDHSHLQEPALCGPHMMFDDBCE	270	
Db	182	EEDRCSHSHKRLCPIDMLMDSNKKCVLQENPLAGTGDHSHLQEPALCGPHMMFDDBCE	241	
QY	271	CYCKTPCPCPDILQHRKNCSEFECKESLFECCCKRHLFPHDTSCSEDRCPFHRRPCASGKT	330	
Db	242	CYCKTPCPCPDILQHRKNCSEFECKESLFECCCKRHLFPHDTSCSEDRCPFHRRPCASGKT	301	

RESULT 2
US-08-915-795-3
Sequence 3, Application US/08915795
Patent No. 6235713
GENERAL INFORMATION:
APPLICANT: Marc G. ACHEN
APPLICANT: Andrew F. WILKS
APPLICANT: Steven A. STACKER
APPLICANT: Karl ALITALO
TITLE OF INVENTION: GROWTH FACTOR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
STREET: 1200 G Street, NW, Suite 700
CITY: Washington
STATE: DC
COUNTRY: United States of America
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,795
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/42983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
TELEX: N/A
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 325 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE: Human Breast
Tissue Type: Human Breast
US-08-915-795-3

RESULT 3
 US-08-915-795-8
 : Sequence 8, Application US/08915795
 : Patent No. 6235713
 : GENERAL INFORMATION:
 : APPLICANT: Marc G. ACHEN
 : APPLICANT: Andrew F. WILKS
 : APPLICANT: Steven A. STACKER
 : APPLICANT: Karl ALITALO
 : TITLE OF INVENTION: GROWTH FACTOR
 : NUMBER OF SEQUENCES: 11
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
 : STREET: 1200 G Street, NW, Suite 700
 : CITY: Washington
 : STATE: DC
 : COUNTRY: United States of America
 : ZIP: 20005
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/915,795
 : FILING DATE:
 : CLASSIFICATION: 536
 : ATTORNEY/AGENT INFORMATION:
 : NAME: EVANS, Joseph D.
 : REGISTRATION NUMBER: 26,269
 : REFERENCE/DOCKET NUMBER: 1064/42983
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (202) 628-8800
 : TELEFAX: (202) 628-8844
 : TELEX: N/A
 : INFORMATION FOR SEQ ID NO: 8:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 358 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein


```
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-999-811-2

Query Match      35.9%; Score 704.5; DB 2; Length 419;
Best Local Similarity 38.9%; Pred. No. 1.1e-56;
Matches 140; Conservative 61; Mismatches 88; Indels 71; Gaps 11;

QY 41 EEOIRASSLEELRLTRTHSEDMKIMRCRL-----KSTSDMSASHSRSTRFAATFY 94
DB 57 EEOIRSVSSVDELMATLYLYPYMKMKCKQLKRGWQHNRQANLSNR--TEETIKFAAHY 114
QY 95 DIETLKVIDEWMORTQCSPRETCVEVASELKGSTNTEFFKPCVNVFRCGCCNEESLIGM 154
DB 115 NTEILKSIDNEWMKTKQCMREVCIDYKKEFGVATNTEFFKPCVSVYRCGCCNSEGLQCM 174
QY 155 NNTSYISQLEIFISVPLTSVPELVYKVAANTGCKCLPTAP--RHPSIIRSI--QIP 210
DB 175 NNTSYISQLEIFISVPLTSVPELVYKVAANTGCKCLPTAP--RHPSIIRSI--QIP 210
QY 211 EEDRCHSKKLCIPIDMLMSNCKCYLQE-----ENPLAGTED----HSHLOE--- 254
DB 235 Q---COAMANKTCPTNMMNNHICRCLAQEDFMFSSDAGDSDTDFGHDICGPNKDELETIC 291
QY 255 -----PALGDPH-----MMFEDRCRCVCKTPCPKDL 281
DB 292 QCVCRAGLRPASGCPHKELDRNSCQCVCNKKLFPSQCGANREDENTCCQVCYKRCPRNQ 351
QY 282 IOHPKNCSECKESLETCCKQKHKLPHPTGSCEDRCRPHTRPCASGKTACAKHCRFPKE 341
DB 352 PLNPGKAC-ECTESPCKLKGKKFHHOTGSC-----YRRPCTNRORACEPGFSYSEE 404

RESULT 6
US-09-042-105-2
Sequence 2, Application US/09042105
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSWEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042.105
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550
FILING DATE: 8-MAR-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
```

```
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TO BE ASSIGNED
FILING DATE: 24-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: ERIC K. STEFFE
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-042-105-2

Query Match      35.9%; Score 704.5; DB 3; Length 419;
Best Local Similarity 38.9%; Pred. No. 1.1e-56;
Matches 140; Conservative 61; Mismatches 88; Indels 71; Gaps 11;

QY 41 EEOIRASSLEELRLTRTHSEDMKIMRCRL-----KSTSDMSASHSRSTRFAATFY 94
DB 57 EEOIRSVSSVDELMATLYLYPYMKMKCKQLKRGWQHNRQANLSNR--TEETIKFAAHY 114
QY 95 DIETLKVIDEWMORTQCSPRETCVEVASELKGSTNTEFFKPCVNVFRCGCCNEESLIGM 154
DB 115 NTEILKSIDNEWMKTKQCMREVCIDYKKEFGVATNTEFFKPCVSVYRCGCCNSEGLQCM 174
QY 155 NNTSYISQLEIFISVPLTSVPELVYKVAANTGCKCLPTAP--RHPSIIRSI--QIP 210
DB 175 NNTSYISQLEIFISVPLTSVPELVYKVAANTGCKCLPTAP--RHPSIIRSI--QIP 210
QY 211 EEDRCHSKKLCIPIDMLMSNCKCYLQE-----ENPLAGTED----HSHLOE--- 254
DB 235 Q---COAMANKTCPTNMMNNHICRCLAQEDFMFSSDAGDSDTDFGHDICGPNKDELETIC 291
QY 255 -----PALGDPH-----MMFEDRCRCVCKTPCPKDL 281
DB 292 QCVCRAGLRPASGCPHKELDRNSCQCVCNKKLFPSQCGANREDENTCCQVCYKRCPRNQ 351
QY 282 IOHPKNCSECKESLETCCKQKHKLPHPTGSCEDRCRPHTRPCASGKTACAKHCRFPKE 341
DB 352 PLNPGKAC-ECTESPCKLKGKKFHHOTGSC-----YRRPCTNRORACEPGFSYSEE 404

RESULT 7
US-09-042-105-18
Sequence 18, Application US/09042105
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSWEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
```


Db 292 QCVCRAGLRPASCGRHKEIDRNSCCVCKNKLFPSSCGANREEDENTCOCVKRGTCPRMQ 351
QY 282 IQHPKNSCECKESLETCCQKHKLFPDPTSCGEDRCPEHTRPCASGKTACAKHCRPKX 341
Db 352 PLNPGKAC-ECTESPCKLKKRKHOTCSC-----YRRPCTNRKACPEGFSTSEE 404
RESULT 11
US-08-795-430-11
Sequence 11, Application US/08795430
Patent No. 6130071
GENERAL INFORMATION:
APPLICANT: Aittalo, Karl
APPLICANT: Joukov, Vladimir
TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
TITLE OF INVENTION: Protein and Gene, Mutants Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,430
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FI96/00427
FILING DATE: 01-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/671,573
FILING DATE: 28-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/601,132
FILING DATE: 14-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/585,895
FILING DATE: 12-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/510,133
FILING DATE: 01-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,011
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28967/33691
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 415 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-795-430-11
Query Match 35.5%; Score 696; DB 4; Length 415;
Best Local Similarity 38.6%; Pred. No. 6.4e-56;
Matches 140; Conservative 57; Mismatches 96; Indels 70; Gaps 11;
QY 27 GPVKRSSOSTLERSQOIRASSLEELRLTHSEDKLWRCRLKLSFTSMDSRSASHRS 86

Db 46 GEVKAEEGKDL---EQILRSVSSVDELMASVLPDYWMYKCOLRKGGMOQPTINTRTGDS 102
QY 87 TRRAATFYDIETLKVIDEEMQRTQCSPREFCVVASLGGSTMTFFRPVANNFRGGCC 146
Db 103 VKPAAHYNTIELKSIDNEMWRTQCMPEVCIQVKEFGAATNFFPVPVSVYRGCGCC 162
QY 147 NEESLQCMNTSTSVYSKOLEISVPLTSVPELVPVKVANNATGCKLPTAP--RHPSIIR 204
Db 163 NSRGLQCMNTSTGYLSKTLPEITVPLSGPKPTIISANTSCRMKSLVYQVHSIIR 222
QY 205 RSI--QIPEEDRCSHKKLIDMLDSNKKCVLQ-----EENPLAGTED---H 249
Db 223 RSLPATLPO---QOANKTCPTNYVMNNYMCRLAODPFIYSNVEDDSTNGFHDVCGPN 279
QY 250 SHQF-----PALGPH-----NMFEDRCRCVC 273
Db 280 KEIDEDTCQCVCKGULPSSCGPHKEIDRNSCCVCKNKLFPNSCGANREEDENTCOCVC 339
QY 274 KTECPDILQHPKNSCECKESLETCCQKHKLFPDPTSCGEDRCPEHTRPCASGKTACA 333
Db 340 KRTCPRNOPLNPGKAC-ECTENTQCKELKGRKHOTCSC-----YRRPCANR---L 388
QY 334 KHC 336
Db 389 KHC 391

RESULT 12
US-08-795-430-13
Sequence 13, Application US/08795430
Patent No. 6130071
GENERAL INFORMATION:
APPLICANT: Aittalo, Karl
APPLICANT: Joukov, Vladimir
TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
TITLE OF INVENTION: Protein and Gene, Mutants Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,430
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FI96/00427
FILING DATE: 01-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/671,573
FILING DATE: 28-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/601,132
FILING DATE: 14-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/585,895
FILING DATE: 12-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/510,133
FILING DATE: 01-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,011
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:48:41 ; Search time 78.16 Seconds

(without alignments)
345.008 Million cell updates/sec

Title: US-09-427-657-4

Perfect score: 1963

Sequence: 1 MYREVVVNVFEMLYLVQV.....HCRPKKRAAGPSRKNP 354

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	704.5	35.9	419	2	S69207
2	200.5	10.2	232	2	A41551
3	200	10.2	190	2	B40080
4	198	10.1	190	2	S52130
5	198	10.1	190	2	B44881
6	198	10.1	214	2	A44881
7	194	9.9	190	2	A35987
8	181.5	9.2	1700	2	S08167
9	176.5	9.0	188	2	JC4680
10	167.5	8.5	146	2	S57956
11	166.5	8.5	120	2	A33787
12	164	8.4	148	2	D49530
13	160.5	8.2	245	1	TYWTS
14	160.5	8.2	149	2	A41236
15	158	8.0	158	2	A56125
16	147.5	7.5	207	2	JC4679
17	145	7.4	133	2	B49530
18	145	7.4	241	1	PFHUG2
19	139.5	7.1	1187	2	T18355
20	135	6.9	241	1	PFHUG2
21	133.5	6.8	225	2	S25097
22	128	6.5	748	2	B28964
23	125.5	6.4	196	2	B28964
24	125.5	6.4	211	1	PFHUG1
25	125	6.4	2946	2	T15840
26	122	6.2	226	1	TYWTS
27	120.5	6.1	160	2	JC0542
28	120	6.1	200	2	T15551
29	120	6.1	215	2	S08220

30	120	6.1	226	2	T51550	platelet-derived g
31	119.5	6.1	1106	2	T44598	hypothetical prote
32	118	6.0	965	2	S62935	hypothetical prote
33	116	5.9	1548	2	S34583	serine proteinase
34	115.5	5.9	1964	2	T09059	notch4 - mouse
35	114.5	5.8	1287	2	A41685	SIL protein - huma
36	114.5	5.8	5376	2	T42215	zonadhesin - mouse
37	113	5.8	197	2	S25096	platelet-derived g
38	113	5.8	846	2	A30889	integrin beta chai
39	112.5	5.7	2195	2	T34264	hypothetical prote
40	112	5.7	3635	2	T10053	laminin alpha 5 ch
41	111.5	5.7	336	2	D69074	polyferredoxin 4x2
42	111.5	5.7	2219	2	T27684	hypothetical prote
43	111	5.7	198	2	J50735	platelet-derived g
44	111	5.7	271	2	A25669	PDGF-related trans
45	111	5.7	2201	2	A32160	tenascin-C - human

ALIGNMENTS

RESULT 1
S69207
vascular endothelial growth factor C precursor - human
N:Alternate names: FLT4 ligand DHM
C:Species: Homo sapiens (man)
C>Date: 27-Apr-1996 #sequence.revision 01-Nov-1996 #text.change 08-Oct-1999
C:Accession: S69207; S61795; S71443; S69208; G02659
R:Joukov, V.; Pajusola, K.; Kaipainen, A.; Chillov, D.; Lahtinen, I.; Kukk, E.; Saksel
EMBO J. 15, 1751, 1996
A>Title: Corrigendum: A novel vascular endothelial growth factor, VEGF-C, is a ligand
A:Reference number: S69207; MUID:96203094
A:Accession: S69207
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-419 <J00>
A:Cross-references: EMBL:X94216; NID:g1177488; PIDD:CA63907.1; PTD:e221096; PTD:g118
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1995
A>Note: only a part of the translation is shown
A>Note: this is a revision to the sequence from reference S61795
R:Joukov, V.; Pajusola, K.; Kaipainen, A.; Chillov, D.; Lahtinen, I.; Kukk, E.; Saksel
EMBO J. 15, 290-298, 1996
A>Title: A novel vascular endothelial growth factor, VEGF-C, is a ligand for the Flt4
A:Reference number: S61795; MUID:96178224
A:Accession: S61795
A>Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 70-419 <J001>
A>Note: this sequence has been revised in reference S69207
A:Accession: S71443
A:Molecule type: protein
A:Residues: 1-104-120 <J002>
R:Lee, J.; Gray, A.; Yuan, J.; Luo, S.M.; Avraham, H.; Wood, W.I.
submitted to the EMBL Data Library, December 1995
A:Description: Vascular endothelial growth factor related protein (VRP): A ligand and
A:Reference number: S69208
A:Accession: S69208
A:Molecule type: mRNA
A:Residues: 1-419 <LEB>
A:Cross-references: EMBL:U43142; NID:g1150988; PIDD:AAA65214.1; PTD:g1150989
R:Morris, J.C.
submitted to the EMBL Data Library, May 1996
A:Reference number: H01557
A:Accession: G02659
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-419 <MOR>
A:Cross-references: EMBL:U58111; NID:g1373426; PIDD:AA802909.1; PTD:g1373427
C:Genetics:
A:Gene: GDB:VEGFC; VRP
A:Cross-references: GDB:3890883; OMIM:601528
F.1-12/Domain: signal sequence #status predicted <SIG>
F.13-102/Domain: propeptide #status predicted <PRO>

RESULT 3

B40080

Vascular endothelial growth factor precursor (version 2) - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 30-Jun-1992 #sequence.revision 30-Jun-1992 #text_change 05-Nov-1999

C:Accession: B40080; B33787; A33255

R:Leung, D.W.; Cachianes, G.; Kiang, W.J.; Goeddel, D.V.; Ferrara, N.

Science 246, 1306-1309, 1989

A:Title: Vascular endothelial growth factor is a secreted angiogenic mitogen.

A:Reference number: A40080; MUID:90069608

A:Accession: B40080

A:Molecule type: mRNA

A:Residues: 1-190 <LEU>

A:Cross-references: GB:M32976; NID:q163006; PIDN:AAA30502.1; PID:q163007

R:Fischer, E.; Gospodarowicz, D.; Mitchell, R.; Silva, M.; Schilling, J.; Lau, K.; Crist

Biochem. Biophys. Res. Commun. 165, 1198-1206, 1989

A:Title: Vascular endothelial growth factor: a new member of the platelet-derived growth

A:Reference number: A33787; MUID:90121225

A:Accession: B33787

A:Molecule type: mRNA

A:Residues: 27-190 <TIS>

A:Cross-references: GB:M31836; NID:q163808; PIDN:AAA30804.1; PID:q163809

R:Ferrara, N.; Henzel, W.J.

Biochem. Biophys. Res. Commun. 161, 851-858, 1989

A:Title: Pituitary follicular cells secrete a novel heparin-binding growth factor specific

A:Reference number: A33255; MUID:89286596

A:Accession: A33255

A:Molecule type: protein

A:Residues: 27-31 <PER>

C:Keywords: alternative splicing; glycoprotein

F:1-26/Domain: signal sequence #status predicted <SIG>

F:27-190/Product: vascular endothelial growth factor #status predicted <MAT>

F:100/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.2%; Score 200; DB 2; Length 190;

Best Local Similarity 25.4%; Pred. No. 5; 9e-08;

Matches 57; Conservative 21; Mismatches 66; Indels 80; Gaps 8;

97 ETLKVIDEENQROGSPRECEVASELKGSTMTFFKPPCVNVRGCGCCNEESLGMNT 156

38 EVKRFMD-VYGRSCRTETLVDFIDFQYDEIEYIRKPCVPLMRGCGCCNDGLCVPT 96

157 STSYISKOLEFISVPLTSVPELVKVNHTGCKCLPTARHPYSIIRSIQPEEDRCS 216

97 EEFNTIMQIMRIK-----PHSGQH-----IGKMSFLQ 123

217 HSKKLCIDMLMDSNCKC-----VLOENPLAGTEDHSHLOEPALCGPHMFDEDRCEC 271

124 H-----NKCCEPRKKRKAQENP-----CGP----- 144

272 VCKTPCKDLIOHPKNSCECKESLETCCQKHKLPHPTCSCE 315

145 -GSERRKHLFVQDPQCKC-SCKNIDSRCAROLELNERCQCD 186

RESULT 4

S52130

Vascular endothelial growth factor - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 14-Jul-1995 #sequence.revision 21-Jul-1995 #text_change 05-Nov-1999

C:Accession: S52130

R:Sharma, H.S.; Tang, Z.H.; Gho, B.C.G.; Verdouw, P.D.

Biochim. Biophys. Acta 1260, 235-238, 1995

A:Title: Nucleotide sequence and expression of the porcine vascular endothelial growth

A:Reference number: S52130; MUID:95143284

A:Accession: S52130

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-190 <SHA>

A:Cross-references: GB:X81380; NID:9587559; PIDN:CAA57143.1; PID:9587560

Query Match 10.1%; Score 198; DB 2; Length 190;

Best Local Similarity 24.2%; Pred. No. 8; 3e-08;

Matches 54; Conservative 24; Mismatches 67; Indels 78; Gaps 8;

97 ETLKVIDEENQROGSPRECEVASELKGSTMTFFKPPCVNVRGCGCCNEESLGMNT 156

38 EVKRFMD-VYGRSCRTETLVDFIDFQYDEIEYIRKPCVPLMRGCGCCNDGLCVPT 96

157 STSYISKOLEFISVPLTSVPELVKVNHTGCKCLPTARHPYSIIRSIQPEEDRCS 212

97 EEFNTIMQIMRIK-----PHSGQH-----IGKMSFLQ 133

213 DRCSHKKLCIDMLMDSNCKC-----VLOENPLAGTEDHSHLOEPALCGPHMFDEDRCEC 272

134 DRA-----KQENP-----CGP----- 144

273 CKTPCKDLIOHPKNSCECKESLETCCQKHKLPHPTCSCE 315

145 GSERRKHLFVQDPQCKC-SCKNIDSRCAROLELNERCQCD 186

RESULT 5

B44881

Vascular endothelial growth factor-1 precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 03-Feb-1994 #sequence.revision 03-Feb-1994 #text_change 05-Nov-1999

C:Accession: B44881; A43351; A61029

R:Breier, G.; Albrecht, U.; Storrer, S.; Risau, W.

Development 114, 521-532, 1992

A:Title: Expression of vascular endothelial growth factor during embryonic angiogenesis

A:Reference number: A44881; MUID:92274860

A:Accession: B44881

A:Molecule type: mRNA

A:Residues: 1-190

A:Cross-references: GB:S38083; NID:9249858; PIDN:AAB22253.1; PID:9249859

A:Experimental source: embryo

A:Note: sequence extracted from NCBI backbone (NCBIN:107622; NCBI:107623)

R:Claffey, K.P.; Wilkison, W.O.; Spiegelman, B.M.

J. Biol. Chem. 267, 16317-16322, 1992

A:Title: Vascular endothelial growth factor. Regulation by cell differentiation and a

A:Reference number: A43351; MUID:92353593

A:Accession: A43351

A:Molecule type: mRNA

A:Residues: 1-116/ER,119-190 <CIA>

A:Cross-references: GB:M95200; NID:9202350; PIDN:AAA40547.1; PID:9202351

A:Note: sequence extracted from NCBI backbone (NCBIN:110653; NCBI:110655)

R:Rosenthal, R.A.; Megyesi, J.F.; Henzel, W.J.; Ferrara, N.; Folkman, J.

Growth Factors 4, 53-59, 1990

A:Title: Conditioned medium from mouse sarcoma 180 cells contains vascular endothelial

A:Reference number: A61029; MUID:91197543

A:Accession: A61029

A:Molecule type: protein

A:Residues: 27-38 <ROS>

C:Keywords: alternative splicing; angiogenesis; dimer; disulfide bond; glycoprotein;

Query Match 10.1%; Score 198; DB 2; Length 190;

Best Local Similarity 22.6%; Pred. No. 8; 3e-08;

Matches 60; Conservative 30; Mismatches 78; Indels 98; Gaps 10;

62 WKLMRCRLT-----KSTMSDRSASHSTRFAATFYIEITLKVIDEENQROGSP 113

7 VVHMTLALLLYLHAKWSQAPRTTEGEQSH-----EVYKFM-D-VYGRSCRP 53

114 RETCEVASELKGSTMTFFKPPCVNVRGCGCCNEESLGMNTSYISKOLEFISVPLT 173

54 IETLVDFIDFQYDEIEYIRKPCVPLMRGCGCCNDGLCVPTSSNTIMQIMRIK-PHQ 112

174 SVPELVKVNHTGCKCLPTARHPYSIIRSIQPEEDRCSHKKLCIDMLMDSNCKC 233

113 S-QHIGEMSFLOHSRCECR-----PKDKRK----- 137

234 KCVLOENPLAGTEDHSHLOEPALCGPHMFDEDRCECCKTPCKD-----LIOHPKNS 289


```

OY 210 PEE-----DRCSSKTL-----CPIDM 227
      ||::||
Db 1180 KRKPCKOCBGODMNNHOCGCCPPRPPPCSNOKYSNVSCGCGNPKPKKGCNGNOI 1239
OY 228 WDSNCKCYLQF--ENP-----LACTEDH 249
      ||::||
Db 1240 WCDNTCRVCYPRKMEKPRADNCKTKMMNDMEQCVCRCGCEGCKGVMKMNANTSCCECP 1299
OY 250 SHLOPALTGPMMDEEDRCCEVCYKTP-----CPKDLIDHPKNCSECEKSELETCCQKH 304
      ||::||
Db 1300 ADKAPASGDKKSMNDSCGCCCKMKPCGCGPPNOMNENKDC---ECKSATGNCPCAG 1356
OY 305 KLEHPTGCEDECRPHFRPCASGKTACAKHCR--PRPKR 343
Db 1357 QTMNSOTCQCS--CP-ATGKCTGAQWVCSKACVCPCPAOK 1394

RESULT 9
JC4680
vascular endothelial growth factor-related factor 167 precursor - mouse
M:Alternate names: VRF 167 protein
C:Species: Mus musculus (house mouse)
C:Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 05-Nov-1999
C:Accession: J04680
R:Towson, S.; Lagercrantz, J.; Glimond, S.; Sillins, G.; Nordenskjold, M.; Weber, G.;
Biochem. Biophys. Res. Commun. 220, 922-928, 1996
A:Title: Characterization of the murine VEGF-related factor gene.
A:Reference number: J04679; MUID:96183052
A:Accession: J04680
A:Molecule type: mRNA
A:Residues: 1-188 <TOM>
A:Cross-references: GB:U43837; NID:91314335; PIDN:AAC5253.1; PID:91314336
C:Comment: This factor is a mitogen, that is selective for endothelial cells, and belong
ar endothelial growth factors 167 and VEGF 186.
C:Genetics:
A:Gene: vrf
A:Map position: 19
A:Introns: 137/2
F:1-21/Domain: Signal sequence #status predicted <SIG>
F:22-188/Product: vascular endothelial growth factor-related factor #status predicted <M
Matches 49; Conservative 22; Mismatches 70; Indels 69; Gaps 7;

Query Match 9.0%; Score 176.5; DB 2: Length 188;
Best Local Similarity 23.3%; Pred. No. 3.2e-06;
Matches 49; Conservative 22; Mismatches 70; Indels 69; Gaps 7;

OY 106 WORTQCSPRETCVAVASELGTSTNTEFFKPPCVAVFRCGCCNDESLICMNTSTYSISKOL 165
      ||::||
Db 42 YARATQCPREVVYVPLSMELMGVNVKQLVPSCVYQRCGCCPDGEGCVPTGOHGYRMQI 101
OY 166 FELSVPLTVPELVKPVKVNHTGCKCLPTAPRHPSIIRSIQIPEDRCSHKIKCPID 225
      ||::||
Db 102 LMIOYPSOLGEM---SLEHSCCECRPKR-----KSAVVKPCSPRI---LCP-- 143
OY 226 MLWDSNCKCYLQENPLAGTEDHSHLOPALTGPMMDEEDRCCEVCYKTPCKDIDHP 285
      ||::||
Db 144 -----PCTQRORRP-----DP 154
OY 286 KNCSECEKSELETCCQKHKL-FHPDTCSC 314
      ||::||
Db 155 RTRCRG-RCRRRRFLHCGRGLINPDTGRC 183

RESULT 10
S57956
ovine vascular endothelial growth factor - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 05-Nov-1999
C:Accession: S57956
R:Redmer, D.A.; Dai, Y.; Li, J.; Jones, S.C.; Moor, R.M.
submitted to the EMBL Data Library, July 1995
A:Reference number: S57956
```

```

A:Accession: S57956
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-146 <RED>
A:Cross-References: EMBL:X89506; NID:9899350; PIDN:CAA61677.1; PID:9899351

Query Match 8.5%; Score 167.5; DB 2: Length 146;
Best Local Similarity 38.1%; Pred. No. 1.2e-05;
Matches 37; Conservative 14; Mismatches 43; Indels 3; Gaps 3;

OY 97 ETLKVIDEEMORTQCSPRETCVAVASELGTSTNTEFFKPPCVAVFRCGCCNDESLICMNT 156
      ||::||
Db 38 EYKKEMD-VYQSFRCPIETLVDFQBYDEIEFTFKPCVPLMRGCGCNDSELCVPT 96
OY 157 STYSISKQLEISVPLTVPELVKPVKVNHTGCKCLP 193
      ||::||
Db 97 EEFNITMQLMRK-PIQS-OHIGEMSFLOHNKCECRP 131

RESULT 11
A33787
vascular endothelial growth factor (version 1) - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 16-Mar-1990 #sequence_revision 16-Mar-1990 #text_change 05-Nov-1999
C:Accession: A33787
R:Ritscher, E.; Gospodarowicz, D.; Mitchell, R.; Silva, M.; Schilling, J.; Lau, K.; Cr
Biochem. Biophys. Res. Commun. 165, 1198-1206, 1989
A:Title: Vascular endothelial growth factor: a new member of the platelet-derived gro
A:Reference number: A33787; MUID:90121225
A:Accession: A33787
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-120 <TIS>
A:Cross-references: GB:M33750; NID:9163810; PIDN:AAA30805.1; PID:9163811
C:Keywords: alternative splicing

Query Match 8.5%; Score 166.5; DB 2: Length 120;
Best Local Similarity 38.1%; Pred. No. 1.2e-05;
Matches 37; Conservative 14; Mismatches 43; Indels 3; Gaps 3;

OY 97 ETLKVIDEEMORTQCSPRETCVAVASELGTSTNTEFFKPPCVAVFRCGCCNDESLICMNT 156
      ||::||
Db 12 EYKKEMD-VYQSFRCPIETLVDFQBYDEIEFTFKPCVPLMRGCGCNDSELCVPT 70
OY 157 STYSISKQLEISVPLTVPELVKPVKVNHTGCKCLP 193
      ||::||
Db 71 EEFNITMQLMRK-PIQS-OHIGEMSFLOHNKCECRP 105

RESULT 12
D49530
16k vascular endothelial growth factor homolog A2R - Orf virus
C:Species: Orf virus
C:Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
C:Accession: D49530
R:Lytelle, D.J.; Fraser, K.M.; Fleming, S.B.; Mercer, A.A.; Robinson, A.J.
J. Virol. 68, 84-92, 1994
A:Title: Homologs of vascular endothelial growth factor are encoded by the poxvirus o
A:Reference number: A49530; MUID:94076465
A:Contents: N27
A:Accession: D49530
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-148 <LYT>
A:Cross-references: GB:S67522; NID:9456900; PIDN:AAB29223.1; PID:9456902
A:Note: sequence extracted from NCBI backbone (NCBIN:141422, NCBI:P:141426)

Query Match 8.4%; Score 164; DB 2: Length 148;
Best Local Similarity 32.8%; Pred. No. 2.2e-05;
Matches 38; Conservative 12; Mismatches 40; Indels 26; Gaps 4;
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:51:46 ; Search time 42.92 Seconds

(without alignments)
282.536 Million cell updates/sec

Title: US-09-427-657-4

Perfect score: 1963

Sequence: 1 MYREWVYVNVEMLYQLVQ.....HCRFPKRAAGCPHSRKNP 354

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	704.5	35.9	419	1	VEGC_HUMAN
2	696	35.5	415	1	VEGC_MOUSE
3	200	10.2	190	1	VEGF_BOVIN
4	198	10.1	190	1	VEGF_PIG
5	198	10.1	214	1	VEGF_MOUSE
6	196	10.0	215	1	VEGF_HUMAN
7	194	9.9	190	1	VEGF_RAT
8	185	9.4	164	1	VEGF_CAVPO
9	181.5	9.2	1700	1	BAR3_CHITE
10	176.5	9.0	188	1	VEGB_MOUSE
11	175.5	8.9	188	1	VEGB_HUMAN
12	167.5	8.5	146	1	VEGF_SHEEP
13	164	8.4	148	1	VEGH_OREN7
14	161	8.2	245	1	PDGB_FELCA
15	159.5	8.1	170	1	PLGF_HUMAN
16	155.5	7.9	216	1	VEGF_CHICK
17	155	7.9	158	1	PLGF_MOUSE
18	148	7.5	133	1	VEGH_OREN2
19	145	7.4	241	1	PDGB_HUMAN
20	135	6.9	241	1	PDGB_MOUSE
21	134.5	6.9	241	1	PDGB_SHEEP
22	133.5	6.8	225	1	PDGB_RAT
23	125.5	6.4	211	1	PDGA_HUMAN
24	122	6.2	226	1	TIS15_SMSAV
25	120	6.1	226	1	PDGA_XENLA
26	118	6.0	965	1	YNC3_YEAST
27	116	5.9	1877	1	PKK5_MOUSE
28	115.5	5.9	1964	1	NTC4_MOUSE
29	113	5.8	204	1	PDGA_RAT
30	113	5.8	846	1	ITBX_DROME
31	112	5.7	3635	1	LM45_MOUSE
32	111	5.7	213	1	PDGA_RABIT
33	111	5.7	2201	1	TENA_HUMAN

34	109	5.6	677	1	SP87_DICTI
35	109	5.6	1104	1	NEX1_HUMAN
36	109	5.6	2871	1	FBX1_HUMAN
37	108	5.5	211	1	PDGA_MOUSE
38	108	5.5	1426	1	EGFR_DROME
39	108	5.5	4544	1	LRP1_HUMAN
40	107.5	5.5	1178	1	TSP2_CHICK
41	107	5.5	1680	1	FUR2_DROME
42	107	5.5	2907	1	FBX2_MOUSE
43	106	5.4	1808	1	TENA_CHICK
44	105.5	5.4	1696	1	PKC5_BRACL
45	105	5.3	191	1	WAP_MACEU

ALIGNMENTS

RESULT	ID	VEGC_HUMAN	STANDARD	PRT	419 AA.
AC	P49767	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	01-OCT-2000 (Rel. 40, Last annotation update)				
DE	VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR (VEGF-C) (VASCULAR				
DE	ENDOTHELIAL GROWTH FACTOR RELATED PROTEIN) (VRF) (FLT4 LIGAND) (FLT4-				
DE	L).				
GN	VEGFC.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 103-120.				
RX	MEDLINE=96178224; PubMed=8617204;				
RA	Joukov V., Patjusola K., Kaipainen A., Chillov D., Lahtinen I., Kuk E.,				
RA	Saksela O., Kalkkinen N., Alitalo K.;				
RT	"A novel vascular endothelial growth factor VEGF-C, is a ligand for				
RT	the Flt4 (VEGFR-3) and KDR (VEGFR-2) receptor tyrosine kinases.";				
RL	EMBO J. 15:290-298(1996).				
RN	[2]				
RP	ERRATUM.				
RX	MEDLINE=96203094; PubMed=8612600;				
RA	Joukov V., Patjusola K., Kaipainen A., Chillov D., Lahtinen I., Kuk E.,				
RA	Saksela O., Kalkkinen N., Alitalo K.;				
RL	EMBO J. 15:1751-1751(1996).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=96312526; PubMed=8700872;				
RA	Lee J., Gray A., Yuan J., Luo S.-M., Avraham H., Wood W.I.;				
RT	"Vascular endothelial growth factor-related protein: a ligand and				
RT	specific activator of the tyrosine kinase receptor Flt4.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 93:1988-1992(1996).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	Fitz L., Morris J.C., Towler P.S., Long A.J., Greco R.,				
RA	Burgess P., Giannotti J., Charletta A., Hennessey D., Kovacic S.,				
RA	Fitzgerald M., Scaltireto H., Welch N., Neben S., Finnelly H.,				
RA	Zollner R., Wang J., Nickbarg E., Gassaway R., Turner K.,				
RA	Wood C.R.;				
RL	Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.				
CC	- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL				
CC	CELL GROWTH.				
CC	- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.				
CC	- PTM: PROBABLY PROTEOLITICALLY PROCESSED IN THE C-TERMINUS.				
CC	- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				


```

Db 223 RSLPALPDQ---CQANKTCPTNYWNNMYMNCRLAQODFIYFYSNVEDSTNGFHVCGFN 279
Oy 250 SHLOE-----PALGSPH-----IMFEDDCCEVC 273
Db 280 KLEDEDTCCVCCKGGRSSCGPHELDKRDSCQCCVCKNKLFPNSCGANKEEPEENTCQVC 339
Oy 274 KTRCPKDLQHPKNCSECFESLETCCKHKLFPHTDSCEDRCPFHTRPCASKTCA 333
Db 340 KTCPCRNQPLNKGKAC--ECSTENTQCKFLKGGKFFHQTSC-----YRRPCANR-----L 388
Oy 334 KHC 336
Db 389 KHC 391

RESULT 3
VEGF_BOVIN STANDARD: PRT: 190 AA.
AC P15691;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR PRECURSOR (VEGF) (VASCULAR PERMEABILITY FACTOR) (VPF).
GN VEGF.
OC Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Kuminantia; Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-47.
RA MEDLINE=90069608; PubMed=2479986;
RT Leung D.W., Cachianes G., Kuang W.-J., Goeddel D.V., Ferrara N.;
RT "Vascular endothelial growth factor is a secreted angiogenic mitogen.";
RL Science 246:1306-1309(1989).
RN [2]
RP SEQUENCE OF 27-190 FROM N.A.
RA MEDLINE=90121225; PubMed=2610687;
RT Tischer E., Gospodarowicz D., Mitchell R., Silva M., Schilling J., Lau K., Crisp T., Fiddes J.C., Abraham J.A.;
RT "Vascular endothelial growth factor: a new member of the platelet-derived growth factor gene family.";
RL Biochem. Biophys. Res. Commun. 165:1198-1206(1989).
RN [3]
RP SEQUENCE OF 27-31.
RA MEDLINE=89286596; PubMed=2735925;
RT Ferrara N., Henzel W.J.;
RT "Pituitary follicular cells secrete a novel heparin-binding growth factor specific for vascular endothelial cells.";
RL Biochem. Biophys. Res. Commun. 161:851-858(1989).
CC -I- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL CELL GROWTH. INDUCES ENDOTHELIAL PROLIFERATION AND VASCULAR PERMEABILITY.
CC -I- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -I- SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb.sib.ch).
CC -----
CC EMBL, M32976; AAA30502.1; -
CC EMBL, M31836; AAA30804.1; -
CC EMBL, M33750; AAA30805.1; -
CC PIR: A33255; A33255.

```

```

DR PIR: A33787; A33787.
DR PIR: B40080; B40080.
DR HSSP: P15692; 2VGH.
DR InterPro; IPR000072; -.
DR Pfam; PF00341; PDGF_1; 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS50278; PDGF_2; 1.
KM Mitogen; Growth factor; Glycoprotein; Alternative splicing; Signal.
FT SIGNAL 1 26
FT CHAIN 27 190 VASCULAR ENDOTHELIAL GROWTH FACTOR.
FT DISULFID 51 93 BY SIMILARITY.
FT DISULFID 82 127 BY SIMILARITY.
FT DISULFID 86 129 BY SIMILARITY.
FT DISULFID 76 76 INTERCHAIN (BY SIMILARITY).
FT DISULFID 85 85 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 100 100 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARSPPLIC 139 183 MISSING (IN ISOFORM BETA).
FT VARSPPLIC 184 184 R -> K (IN ISOFORM BETA).
SQ SEQUENCE 190 AA; 22310 MW; EDBF903E46E24789.CRC64;

Query Match 10.2%; Score 200; DB 1; Length 190;
Best Local Similarity 25.4%; Pred. No. 4.1e-09;
Matches 57; Conservative 21; Mismatches 66; Indels 80; Gaps 8;

Oy 97 ETLKVIDEEMORTQCSPRETCVEVASSELGKSTNFFKPPCVAVFPGGCCNEESLICMT 156
Db 38 EVAKFMD-VYORSFCRPLETLVDIEFPEDETFKPSVPLMKGGCCNDESLECVPT 96
Oy 157 STSYISKOLFETSVPLTSPVLVPKVNHTGCKCLPTAPRHPYSIIRSIQIPEDRCS 216
Db 97 EEFNTMQIMRK-----PHQSH-----IGMSFLQ 123
Oy 217 HSKRLCPIDMLMDSNKKCC-----VLQENPLAGTSHLSHLPALCGPHMFDEDRCC 271
Db 124 H-----NKCGRPKDKAROENP-----CGP----- 144
Oy 272 VKTPCPKDLQHPKNCSECFESLETCCKHKLFPHTDSCSE 315
Db 145 -CSERRKHLFVDPOTCK-SCKNTDSRCKARQDELNERTCKCD 186

RESULT 4
VEGF_PIG STANDARD: PRT: 190 AA.
AC P49151;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR PRECURSOR (VEGF) (VASCULAR PERMEABILITY FACTOR) (VPF).
GN VEGF.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Heart;
RA MEDLINE=95143284; PubMed=7841203;
RT Sharma H.S., Tang Z.H., Gho B.C.H., Verdouw P.D.;
RT "Nucleotide sequence and expression of the porcine vascular endothelial growth factor.";
RL Biochim. Biophys. Acta 1260:235-238(1995).
CC -I- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL CELL GROWTH. INDUCES ENDOTHELIAL PROLIFERATION AND VASCULAR PERMEABILITY (BY SIMILARITY).
CC -I- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.

```


OY 114 RETCEVASELSTINTFFKPCVNVFRCGCCNESSLICMTSTSYISKOLFELISPLT 173
 DB 54 IETVLVPEEYIEIKRPSVPLMRCAAGCCNDALCEVPTSEENITMQRIR-PHQ 112
 OY 174 SVELVAVKVAHNTGCKCLPTAPRHPYIIRRSIQIPEDRGSHSKLCPIDMLDSNKC 233
 DB 113 S-OHIEMSFLOHSRCECR-----PKKDRTPKPK-----SVRGKKG 149
 OY 234 KCVLOEENPLAGEDSHLQEPALCGPHMFEDRECEVCCKTTPD-----LIQHKNCS 289
 DB 150 QKRRKRSRFSKSVH-----CE-----PSEBRKRLHFVQDPOTCK 185
 OY 290 CFECKSELETCCKKHLFHPDTCSE 315
 DB 186 C-SCKNTDSKCRKARQLEINERTCRCD 210
 RESULT 6
 VEGF HUMAN STANDARD; PRT; 215 AA.
 AC P15692;
 DT 01-APR-1990 (rel. 14, Created)
 DT 01-APR-1990 (rel. 14, Last sequence update)
 DT 15-JUL-1999 (rel. 38, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR PRECURSOR (VEGF) (VASCULAR PERMEABILITY FACTOR) (VPF).
 GN VEGF OR VEGFA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-90069608; PubMed-2479986;
 RX Leung D.W., Cachanes G., Kuang W.-J., Goeddel D.V., Ferrara N.;
 RT "Vascular endothelial growth factor is a secreted angiogenic mitogen.";
 RT Science 246:1306-1309(1989).
 RN [2]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RA MEDLINE-90069609; PubMed-2479987;
 RX Keck P.J., Hauser S.D., Krivt G., Sanzo K., Warren T., Feder J.,
 RA Connolly D.T.;
 RT "Vascular permeability factor, an endothelial cell mitogen related to PDGF.";
 RT Science 246:1309-1312(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-91268072; PubMed-1711045;
 RX Fischer E., Mitchell R., Hartman T., Silva M., Gospodarowicz D.,
 RA Fiddes J.C., Abraham J.A.;
 RT "The human gene for vascular endothelial growth factor. Multiple protein forms are encoded through alternative exon splicing.";
 RT J. Biol. Chem. 266:11947-11954(1991).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-92231879; PubMed-1567395;
 RX Weinzel K., Marne D., Weich H.A.;
 RT "AIDS-associated Kaposi's sarcoma cells in culture express vascular endothelial growth factor.";
 RT Biochem. Biophys. Res. Commun. 183:1167-1174(1992).
 RN [5]
 RP PRELIMINARY SEQUENCE OF 27-36; 43-50 AND 59-81.
 RA MEDLINE-90062112; PubMed-2584205;
 RX Connolly D.T., Olander J.V., Heuvelman D., Nelson R., Monseil R.,
 RA Siegel N., Haymore B.L., Leimgruber R., Feder J.;
 RT "Human vascular permeability factor. Isolation from U937 cells.";
 RT J. Biol. Chem. 264:20017-20024(1989).
 RN [6]
 RP SEQUENCE OF 27-41.
 RA MEDLINE-93145946; PubMed-7678805;
 RX Fiebig B.L., Jaeger B., Schoellmann C., Weinzel K., Wilting J.,
 RA Kochs G., Marne D., Hug H., Weich H.A.;

RT "Synthesis and assembly of functionally active human vascular endothelial growth factor homodimers in insect cells.";
 RT Eur. J. Biochem. 211:19-26(1993).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 34-135.
 RA MEDLINE-97352774; PubMed-9207067;
 RX Muller Y.A., Li B., Christinger H.W., Wells J.A., Cunningham B.C.,
 RA de Vos A.M.;
 RT "Vascular endothelial growth factor: crystal structure and functional mapping of the kinase domain receptor binding site.";
 RT Proc. Natl. Acad. Sci. U.S.A. 94:7192-7197(1997).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (1.93 ANGSTROMS) OF 34-135.
 RA MEDLINE-98035455; PubMed-9351807;
 RX Muller Y.A., Christinger H.W., Keyt B.A., de Vos A.M.;
 RT "The crystal structure of vascular endothelial growth factor (VEGF) refined to 1.93-A resolution: multiple copy flexibility and receptor binding.";
 RT Structure 5:1325-1338(1997).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 39-134.
 RA MEDLINE-99119204; PubMed-9922142;
 RX Wiesmann C., Christinger H.W., Cochran A.G., Cunningham B.C.,
 RA Fairbrother W.J., Keenan C.J., Meng G., de Vos A.M.;
 RT "Crystal structure of the complex between VEGF and a receptor-blocking peptide.";
 RT Biochemistry 37:17765-17772(1998).
 RN [10]
 RP STRUCTURE BY NMR OF 34-135.
 RA MEDLINE-9747915; PubMed-9336848;
 RX Fairbrother W.J., Champe M.A., Christinger H.W., Keyt B.A.,
 RA Starovasnik M.A.;
 RT "The crystal structure of the heparin-binding domain of vascular endothelial growth factor.";
 RT Structure 6:637-648(1998).
 RN [11]
 RP STRUCTURE BY NMR OF 137-215.
 RA MEDLINE-98298440; PubMed-9634701;
 RX Fairbrother W.J., Champe M.A., Christinger H.W., Keyt B.A.,
 RA Starovasnik M.A.;
 RT "Solution structure of the heparin-binding domain of vascular endothelial growth factor.";
 RT Structure 6:637-648(1998).
 RN [12]
 RP FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL CELL GROWTH. INDUCES ENDOTHELIAL PROLIFERATION AND VASCULAR PERMEABILITY.
 CC -I SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -I SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN (BY SIMILARITY).
 CC -I ALTERNATIVE PRODUCTS: FOUR FORMS OF VEGF ARE PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE (VEGF-121, VEGF-165, VEGF-189 AND VEGF-215).
 CC -I SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL, M32977; AAA35789.1; -;
 DR EMBL, M27281; AAA36807.1; -;
 DR EMBL, M63978; AAA36804.1; -;
 DR EMBL, M63971; AAA36804.1; JOINED.
 DR EMBL, M63972; AAA36804.1; JOINED.
 DR EMBL, M63973; AAA36804.1; JOINED.
 DR EMBL, M63974; AAA36804.1; JOINED.
 DR EMBL, M63975; AAA36804.1; JOINED.
 DR EMBL, M63976; AAA36804.1; JOINED.
 DR EMBL, M63977; AAA36804.1; JOINED.


```

DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR (VEGF) (VASCULAR PERMEABILITY
DE FACTOR) (VEGF).
GN VEGF.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.
OX NCBI_TaxId=10141;
RN [1]
RP SEQUENCE FROM N.A.
RA Berris B.;
RL Submitted (xxx-1992) to the EMBL/Genbank/DBJ databases.
CC -! FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
CC CELL GROWTH. INDUCES ENDOTHELIAL PROLIFERATION AND VASCULAR
CC PERMEABILITY.
CC -! SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -! SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR
CC TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN (BY
CC SIMILARITY).
CC -! SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M84230; AAA37057.1; -
DR HSSP; P15692; 2VGH.
DR InterPro: IPR000072; -
DR Pfam: PF00341; PDGF_1;
DR PROSITE; PS00249; PDGF_1;
DR PROSITE; PS50278; PDGF_2; 1.
KW Mitogen; Growth factor; Glycoprotein.
FT DISULFID 25 67 BY SIMILARITY.
FT DISULFID 56 101 BY SIMILARITY.
FT DISULFID 60 103 BY SIMILARITY.
FT DISULFID 50 50 INTERCHAIN (BY SIMILARITY).
FT DISULFID 59 59 INTERCHAIN (BY SIMILARITY).
FT CARBOHD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 164 AA; 19330 MW; 9EB86A81A9D5DCA4 CRC64;

Query Match 9.4%; Score 185; DB 1; Length 164;
Best Local Similarity 24.1%; Pred. No. 5.4e-08;
Matches 54; Conservative 23; Mismatches 67; Indels 80; Gaps 8;
OY 97 ETLKVIDEEMORTGSPRETCVEVASELGSNTFFKPPCVNVR-----CGGCNEESLTCMNTSTYSIK 156
DB 12 EEVKEND-VKRSYCRPIELMVDIFOEYFDELEYTFKPSVCLMRGGGCGNDSLECYVF 70
OY 157 STSYISKOLFELISVPLTVPVAVANHTCCCLPTAPRHPYSIIIRSIQIPEEDRCS 216
DB 71 EEFNITMQLRIK-----PHGGQH-----IGKESFLQ 97
OY 217 HSKKLCPLDMLDSNKKC-----VLQENPLAGTEDHSHLOPALCGPHMFFDEDRCC 271
DB 98 HS-----KCECRPKRERAKRQNP-----CGP----- 118
OY 272 VKTPCPKDILQHPKNCSEFCKESLETCCKHKLPHPTSCSE 315
DB 119 -CSERKHLFVODPOTCK-SCRNTDSRCKAROLELNTENTCKCD 160

RESULT 9
BAR3_CHITE
ID BAR3_CHITE STANDARD; PRT: 1700 AA.
AC 00376;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)

```

```

DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE BALBIANI RING PROTEIN 3 PRECURSOR.
GN BR3.
OS Chironomus tentans (Midge).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomidae; Chironomidae; Chironominae; Chironomus.
OX NCBI_TaxId=7153;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-Salivary gland;
RX MEDLINE=90172404; PubMed=1689777;
RA Paulsson G., Lendah U., Galli J., Ericsson C., Mieslander L.;
RT "The Balbiani ring 3 gene in Chironomus tentans has a diverged
RT repetitive structure split by many introns."
RL J. Mol. Biol. 211:331-349(1990).
CC -! FUNCTION: USED BY THE LARVAE TO CONSTRUCT A SUPRAMOLECULAR
CC STRUCTURE, THE LARVAL TUBE. BALBIANI RING PROTEIN 3 COULD PLAY A
CC ROLE AS A TRANSPORT PROTEIN THAT BINDS TO OTHER PROTEINS
CC INTRACELLULARLY AND IN THE GLAND LUMEN IN ORDER TO PREVENT THESE
CC FROM FORMING WATER-INSOLUBLE FIBERS TOO EARLY.
CC -! SUBCELLULAR LOCATION: SECRETED.
CC -! TISSUE SPECIFICITY: SALIVARY GLAND.
CC -! DOMAIN: HAS 82 APPROXIMATE REPEATS OF CYS-X-CYS-X-CYS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52263; CA36506.1; -
DR PIR; S08167; S08167.
DR HSSP; P18055; 2MRB.
DR InterPro: IPR000853; -
DR PRINTS; PRO0876; MTNEMATODE.
KW Repeat; Signal.
FT SIGNAL 1 1700 POTENTIAL.
FT CHAIN ? 1700 BALBIANI RING PROTEIN 3.
SQ SEQUENCE 1700 AA; 186145 MW; 34202B28521B0815 CRC64;

Query Match 9.2%; Score 181.5; DB 1; Length 1700;
Best Local Similarity 19.9%; Pred. No. 1.1e-06;
Matches 68; Conservative 43; Mismatches 105; Indels 125; Gaps 15;
OY 109 TGCSPRETCVEVASELGSNTFFKPPCVNVR-----CGGCNEESLTCMNTSTYSIK 163
DB 1073 TGCSDKOKRIEESKCGCGCTQF-----QCKDGRFWSLBECCGLCDKK--CP-----GK 1119
OY 164 QLEFISVPLTVPVAVANHT-----TGCKCLPTAPRHPYSIIR-----SIQI 209
DB 1120 QVFDKNTCCQCKPNOKPBDTCNGKDFCLDSCCKKNKPRANGCGVGEWNEKQCCBC 1179
OY 210 PEE-----DRGSHSKL-----CPIDML 227
DB 1180 PDKCKPKCGGQGDWNNHOCQCGCPTPAFTCSNOKYSNVSCGCGNPGKPRNGGNGDI 1239
OY 228 WDSNCKCVLQE--ENP-----LAGTEDH 249
DB 1240 WCDNTRCRCCPRNMKRPADNCKTKMMNDMCCQVCKPGCEBGGCKGVMMKNANTSCSEPR 1299
OY 250 SHLOPALCGPHMFEDEDRCEVCYKTP-----CPDILQHPKNCSEFCKESLETCCKQH 304
DB 1300 ADKARPASCGRKSWNDSCSCQCKSKMPCCGCPNQQWNERDC--ECKSATGNCPRAG 1356
OY 305 KLFHPDTCSEDRCPFHTRPCASGKTACKKQR--FPKCKR 343
DB 1357 QTWNSQTCQCS--CP-ATGKGTGAQVWCSSKACKCVCPAOKK 1394

```

```

RESULT 10
VEGF_MOUSE
ID VEGF_MOUSE STANDARD: PRT: 188 AA.
AC P49766;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR B PRECURSOR (VEGF-B) (VASCULAR
GN ENDOTHELIAL GROWTH FACTOR RELATED PROTEIN) (VRF).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=96197355; PubMed=8637916;
RA Olofsson B., Pajusola K., Kaipainen A., von Euler G., Joukov V.,
RA Saksela O., Orpana A., Pettersson R.F., Alltalo K., Eriksson U.;
RT "Vascular endothelial growth factor B, a novel growth factor for
RT endothelial cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:2576-2581(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96183052; PubMed=8607868;
RA Tomson S., Lagercrantz J., Grimmond S., Silins G.,
RA Nordenskjöld M., Weber G., Hayward N.K.;
RT Characterization of the murine VEGF-related factor gene.";
RL Biochem. Biophys. Res. Commun. 220:922-928(1996).
CC -1- FUNCTION: GROWTH FACTOR FOR ENDOTHELIAL CELLS. BINDS HEPARIN.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED. CAN ALSO FORM HETERODIMER
CC WITH VEGF.
CC -1- SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR
CC TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN.
CC -1- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN HEART, BRAIN, KIDNEY
CC AND SKELETAL MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; U48800; AAB06273.1; -
CC EMBL; U43837; AAC52553.1; -
CC HSSP; P15692; 2VGH.
CC MGD; MGI:106199; Vegfb.
CC InterPro: IPR000072; -
CC Pfam: PF00341; PDGF_1; 1.
CC PROSITE; PS00249; PDGF_1; 1.
CC PROSITE; PS50278; PDGF_2; 1.
CC Mitogen; Growth factor; Signal; Heparin-binding.
CC SIGNAL 1 21 POTENTIAL.
CC CHAIN 22 188 VASCULAR ENDOTHELIAL GROWTH FACTOR B.
CC SEQUENCE 188 AA; 21442 MW; D52A055FB955E9CA CRC64;
SO
Query Match 9.0%; Score 176.5; DB 1; Length 188;
Best Local Similarity 23.3%; Pred. No. 2.9e-07;
Matches 49; Conservative 22; Mismatches 70; Indels 69; Gaps 7;
QY 106 WDRFGSPRETCVEVASLGGKSTNFFKPPCVNVFRCGGCCNEESLICMNTSTSYISKOL 165
: | | | | | : | | | | | : | | | | | : | | | | | :
DB 42 YAAATGCPREVVVPLSMELMGVNVKOLVPSVTVGRCGGCCPDGDECVPTGQHVRMGI 101
: | | | | | : | | | | | : | | | | | : | | | | | :
QY 166 FELSVGLTVSPELVLPVKVANHNGCKCLPTAPRHPYIIRSLIDIPEDKCSHKSLCPID 225
: | | | | | : | | | | | : | | | | | : | | | | | :
DB 102 LMQVPSQLGEM--SLEHSGCCRCRPRK-----KESAVKPDSPRI-----LCP-- 143
: | | | | | : | | | | | : | | | | | : | | | | | :

```

```

QY 226 MLMDSNKKCKVLQENPLAGTDSHILQEPALCGPHMFEDRCCEVCCKTPCKDLIQHP 285
: | | | | | : | | | | | : | | | | | : | | | | | :
DB 144 -----PCTQRRQRP-----DP 154
QY 286 KNCSCFECKESLETCCCKHKL-FHPDTCSG 314
: | | | | | : | | | | | : | | | | | : | | | | | :
DB 155 RRCRC-RCRRRRRLHCQGLNELNPDTCRC 183
: | | | | | : | | | | | : | | | | | : | | | | | :
RESULT 11
VEGF_HUMAN
ID VEGF_HUMAN STANDARD: PRT: 188 AA.
AC P49765;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR B PRECURSOR (VEGF-B) (VEGF RELATED
GN FACTOR).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=96197355; PubMed=8637916;
RA Olofsson B., Pajusola K., Kaipainen A., von Euler G., Joukov V.,
RA Saksela O., Orpana A., Pettersson R.F., Alltalo K., Eriksson U.;
RT "Vascular endothelial growth factor B, a novel growth factor for
RT endothelial cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:2576-2581(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97077124; PubMed=8919691;
RA Grimmond S., Lagercrantz J., Drinkwater C., Silins G., Tomson S.,
RA Pollock P., Golley D., Carson E., Rakar S., Nordenskjöld M., Ward L.,
RA Hayward N., Weber G.;
RT "Cloning and characterization of a novel human gene related to
RT vascular endothelial growth factor.";
RL Genome Res. 6:124-131(1996).
CC -1- FUNCTION: GROWTH FACTOR FOR ENDOTHELIAL CELLS. BINDS HEPARIN.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED. CAN ALSO FORM HETERODIMER
CC WITH VEGF.
CC -1- SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR
CC TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXCEPT LIVER.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXCEPT LIVER.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; U48801; AAB06274.1; -
CC EMBL; U43369; AAA91463.1; -
CC HSSP; P15692; IVPF.
CC MIM; 601398; -
CC InterPro: IPR000072; -
CC Pfam: PF00341; PDGF_1; 1.
CC PROSITE; PS00249; PDGF_1; 1.
CC PROSITE; PS50278; PDGF_2; 1.
CC Mitogen; Growth factor; Signal; Heparin-binding.
CC SIGNAL 1 21 POTENTIAL.
CC CHAIN 22 188 VASCULAR ENDOTHELIAL GROWTH FACTOR B.
CC SEQUENCE 188 AA; 21261 MW; F04654D5A3727194 CRC64;
SO
Query Match 8.9%; Score 175.5; DB 1; Length 188;

```

Best Local Similarity 24.08; Pred. No. 3.4e-07;
Matches 52; Conservative 23; Mismatches 73; Indels 69; Gaps 8

100 KVID--EEMQRTQCSPRETCVEVASELGKSTNTFFKPPCVNVERGGCCNEESLICMNTS 157

34 KVVSWIDVYTRATCQPREVVVPLTVELMGTVAKQLVPSCVTVQRCGGCCPDDGLECVPTG 93

158 TSYISKQLEISVPLTSVPELVKVANHTGCKCLPTAPRHPYSII RRSIQIPEEDRCSH 217

94 QHÖVRMÖILMIRYPSSÖLGEM---SLEEHSÖCECRPK-----KDSAVKPDSPR--- 139

218 SKKLCPI DMLWDSNKKCKV LQ EENPLAGT EDHSH LQEPALCGPHMMFDEDRCECVCKTPC 277

```

140 --PLCP-----RCT-----QHQR-----DPRTCRCRCR--- 162

```

2/8 PKDLIQPNCSCECKESLETCQKHLEHPRDJCS 314

103 ----- KKSF LKCYGKGLNPDICK 183

RESULT 12

VEGF_SHEEP	STANDARD;	PRT;	146 AA.
------------	-----------	------	---------

01-OCT-1996 (Rel. 34, Created)

01-NOV-1997 (Rel. 35, Last annotation update)

PERMEABILITY FACTOR). (VPF).
VEGE
N

Chordata: Vertebrata: Euteleostomi: Eukaryota: Metazoa: Ovis aries (Sheep).

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Caprinae; Ovis.

```
NCBI_TaxID=9940;
[1]
N
```

SEQUENCE FROM N.A.
TISSUE=KIDNEY;

Redmer D.A., Dal Y., Li J., Charnock-Jones D.S., Smith S.K.,
MEDLINE=9/11/536; PubMed=69568642;

Characterization and expression of vascular endothelial growth

J. Reprod. Fertil. 108:157-165(1996).

CELL GROWTH. INDUCES ENDOTHELIAL PROLIFERATION AND VASCULAR

-1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.

TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN (BY

FT	DISULFID	76	76	INTERCHAIN (BY SIMILARITY).
FT	DISULFID	85	85	INTERCHAIN (BY SIMILARITY).
FT	CARBOHD	100	100	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	146 AA;	17247 MM;	4E792CB557F91760 CRC64;

Query Match	8.58;	Score 167.5;	DB 1;	Length 146;
-------------	-------	--------------	-------	-------------

Matches	37;	Conservative	14;	Mismatches	43;	Indels	3;	Gaps	3;
---------	-----	--------------	-----	------------	-----	--------	----	------	----

97 ETLKVIDEEMORTQCSPRETCVEVASSELGKSTNTFFKPPCVNVFRGGCCNEESLICMNT 156

Db 38 EVMKFMD-VYQSFRCRIETLVDFQEPDEIEFIFKPSCVPLMRGGCNDSELECVPT 96

QY 157 STSYISKQLEISVPLTSVPELVPVKVANHTGCKLP 193 .

Db 97

RESULT 13

VEGH_ORFN/	STANDARD;	PRT;	148 AA.
ID_VEGH_ORFN7			

AC	P52585;
DT	01-OCT-1996 (Rel. 34, Created)

DT	15-JUL-1999	(Rel. 38, Last annotation update)
DI	01-OCT-1990	(Rel. 34, Last sequence update)

DE VACCOLAN ENDOHELEIN GROWTH FACTOR HOMOLOG PRECURSOR,
GN A2R.

Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

NCBI_TaxID=73495;

RP SEQUENCE FROM N.A.
MEDT-TW-04026 AGE: D-8-10-2 02E4700

RA Lyttle D.J., Fraser K.M., Fleming S.B., Mercer A.A., Robinson A.J.;
"Homologs of vascular endothelial growth factor are encoded by the

RT poxvirus orf virus." ;
PT 1 Vir01 68-84-9371994)

CC - FUNCTION: INDUCES ENDOTHELIAL PROLIFERATION.
CC - SUBUNIT: HOMODIMER. DISULFIDE-LINKED (BY SIMILARITY)
CC

CC -|- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -|-

cc between the Swiss Institute of Bioinformatics and the EMBL Outstation -
cc This Swiss-Prot entry is copyright. It is produced through a collaboration
cc

Query Match	8.43;	Score 164;	DB 1;	Length 148;
Best Local Similarity	32.88;	Pred. No. 2.1e-06;		
Matches 38;	Conservative 12;	Mismatches 40;	Indels 26;	Gaps 4;

OY 105 EMORT-----QCSPRECEVASELSKSTNFFKPCVNVPRGCGCNESLIOMNTSTSY 160
 DB 36 DMRITDKSGCKRDRVYVIGETPESTINQINPCVYKRSGCCNGGQICITAVETRN 95
 OY 161 ISKOLFEISVPLNSV-----PELVPVKVNHTGCKCL-----PTAPRHP 199
 DB 96 TT-----VTVSVTVGSSSGTNGSVSTNLRISVTEHTKDCICGRITTTPTTREP 146

RESULT 14
 PDB_FELCA STANDARD; PRT: 245 AA.
 AC P12919;
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PLATELET-DERIVED GROWTH FACTOR, B CHAIN PRECURSOR (PDGF B-CHAIN)
 DE (PDGFB) (C-SIS) (PDGF-2).
 GN PDGFB OR SIS.
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 OX NCBI_TaxID=9685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87146463; PubMed=3822831;
 RA van den Ouweland A.M.W., van Groningen J.J.M., Schalken J.A.,
 RA van Neck H.W., Bloemers H.P.J., van de Ven W.J.M.;
 RT "Genetic organization of the c-sis transcription unit.";
 RL Nucleic Acids Res. 15:959-970(1987).
 RN [2]
 RP REVISIONS.
 RA van den Ouweland A.M.W.;
 RL Submitted (Nov-1996) to the EMBL/Genbank/DBJ databases.
 CC -I- FUNCTION: PLATELET-DERIVED GROWTH FACTOR IS A POTENT MITOGEN FOR
 CC CELLS OF MESENCHYMAL ORIGIN. BINDING OF THIS GROWTH FACTOR TO ITS
 CC AFFINITY RECEPTOR ELICITS A VARIETY OF CELLULAR RESPONSES. IT IS
 CC RELEASED BY PLATELETS UPON WOUNDING AND CELLS AN IMPORTANT ROLE
 CC IN STIMULATING ADJACENT CELLS TO GROW AND THEREBY HEAL THE WOUND.
 CC -I- SUBUNIT: ANTIPARALLEL DISULFIDE-LINKED DIMER OF NONIDENTICAL (A
 CC AND B) CHAINS. HOMODIMERS OF A AND B CHAINS ARE IMPLICATED IN
 CC TRANSFORMATION PROCESSES.
 CC -I- MISCELLANEOUS: A-A AND B-B, AS WELL AS A-B, DIMERS CAN BIND TO THE
 CC PDGF RECEPTOR.
 CC -I- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X05112; CAA28758.1; ALT_SEQ.
 DR PIR: A26402; TVCTSS.
 DR HSSP: P01127; 1PDG.
 DR InterPro: IPR002402;
 DR InterPro: IPR002400;
 DR Pfam: PF00341; PDGF; 1.
 DR PRINTS: PR00438; GFCYSKNOT.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS50278; PDGF_2; 1.
 KM Mitogen; Growth factor; Proto-oncogene; Platelet; Signal.
 FT SIGNAL 1 20
 FT PROPEP 21 81
 FT CHAIN 82 194
 FT PROPEP 195 245
 FT DISULFID 101 145
 FT DISULFID 134 182
 FT DISULFID 138 184
 FT DISULFID 128 128

FT DISULFID 137 137 INTERCHAIN (BY SIMILARITY).
 SQ SEQUENCE 245 AA; 27787 MW; E7715291D9837512 CRC64;
 Query Match 8.2%; Score 161; DB 1; Length 245;
 Best Local Similarity 29.9%; Pred. No. 6,1e-06;
 Matches 67; Conservative 24; Mismatches 103; Indels 30; Gaps 10;
 OY 1 MYREVVYVAVFMML--YVQLYGSSNEHGPVKRSSQSTLEREQQIRAASSLLELRTH 58
 DB 1 MNRCA---LFUSLCYKLV---SAEGDPIPEELYKML--SDHSIR--SFDDLRLLH 49
 OY 59 SEDWKIMRCRLRKSFST----MDSRSASHRSRFPATFYDIETLKYIDEMORTQCS 113
 DB 50 GSVDEDRALDNLNSTRSGGELESLSGRRSLEAGSPVAPAMIAE-----CKT 103
 OY 114 RETCEVASELSKSTNFFK--PCVNVFRCGCCNBSLIOMNTSTYSIQOLEIS-V 170
 DB 104 RTEVEFVSRLIDRTNANFLWMPCEVQRCSCNNRNWQCRPTQVQLRLVQVKRIEIV 163
 OY 171 PLTSPELVPVKVNHTGCKCLPTAPRHPYSTRSIQIPEEDR 214
 DB 164 RRPVFKKATVLEDLHACKCEVVAAR--VTRSPSSSQSOR 204

RESULT 15
 PDB_HUMAN STANDARD; PRT: 170 AA.
 AC P49763;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PLACENTA GROWTH FACTOR PRECURSOR (PLGF-1/PLGF-2).
 DE PGF OR PLGF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (PLGF-1).
 RC TISSUE=Placenta;
 RX MEDLINE=92021031; PubMed=1924389;
 RA Maglione D., Guerriero G., Vigiiletto G., Dell-I-Boyi P., Persico M.G.;
 RT "Isolation of a human placenta cDNA coding for a protein related to
 RT the vascular permeability factor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9267-9271(1991).
 RN [2]
 RP SEQUENCE FROM N.A. (PLGF-2).
 RC TISSUE=Placenta;
 RX MEDLINE=94198032; PubMed=8148155;
 RA Hauser S., Welch H.A.;
 RT "A heparin-binding form of placenta growth factor (PLGF-2) is
 RT expressed in human umbilical vein endothelial cells and in
 RT placenta.";
 RL Growth Factors 9:259-268(1993).
 RN [3]
 RP SEQUENCE FROM N.A. (PLGF-2).
 RX MEDLINE=93205407; PubMed=7681160;
 RA Maglione D., Guerriero V., Vigiiletto G., Ferraro M.G., Aprelikova O.,
 RA Allitalo K., del Vecchio S., Lei K.-J., Chou J.Y., Persico M.G.;
 RT "Two alternative mRNAs coding for the angiogenic factor, placenta
 RT growth factor (PLGF), are transcribed from a single gene of
 RT chromosome 14.";
 RL Oncogene 8:925-931(1993).
 RN [4]
 RP CHARACTERIZATION AND SEQUENCE OF 19-24.
 RX MEDLINE=95014370; PubMed=7929268;
 RA Park J.E., Chen H.H., Winer J., Houck K.A., Ferrara N.;
 RT "Placenta growth factor. Potentiation of vascular endothelial growth
 RT factor bioactivity, in vitro and in vivo, and high affinity binding
 RT to Flt-1 but not to Flk-1/KDR.";
 RT J. Biol. Chem. 269:25646-25654(1994).
 CC -I- FUNCTION: GROWTH FACTOR OF UNKNOWN FUNCTION. BINDS TO RECEPTOR

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:50:55 ; Search time 128.06 Seconds
(without alignments)
365.735 Million cell updates/sec

Title: US-09-427-657-4
Perfect score: 1963
Sequence: 1 MYREVVVVNFMMLYVOLVQ.....HCRPPEKRAAQQPHSRKNP 354

Scoring table:
BLASTSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP.TREMBL.16:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mhc:*
8: sp.mammal:*
9: sp.phage:*
10: sp.plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1963	100.0	354	4	043915 homo sapien
2	1675	85.3	358	11	P97946 mus musculu
3	1517.5	77.3	326	11	035251 rattus norv
4	692	35.3	418	13	057352
5	682.5	34.8	420	6	09XS50
6	644	32.8	122	6	09GLX1
7	287	14.6	126	11	035757
8	204	10.4	254	4	016889
9	203	10.3	190	6	09GKRO
10	201	10.2	190	6	077643
11	201	10.2	209	4	060720
12	200.5	10.2	232	4	09H1W9
13	199	10.1	190	6	09XS53
14	198.5	10.1	194	13	042572
15	198.5	10.1	1704	5	094446
16	198	10.1	190	6	09GL52
17	198	10.1	190	11	09OX39
18	196	10.0	208	6	09XSF4
19	195	9.9	191	4	075875

20	194	9.9	214	11	09OXG7	09qgx7 rattus norv
21	193	9.8	214	6	09MWV3	09mwv3 canis famli
22	191.5	9.8	142	11	09ERT6	09ert6 mesocricetu
23	191	9.7	214	6	09XSF5	09xsf5 canis famli
24	189	9.6	174	4	09UL23	09ul23 homo sapien
25	181.5	9.2	1698	5	094438	094438 chronomus
26	179.5	9.1	188	6	09XS48	09xs48 bos taurus
27	171.5	8.7	170	11	09JRX7	09jrx7 rattus norv
28	171	8.7	171	4	09H1W8	09h1w8 homo sapien
29	170	8.7	124	6	09GKRO	09gkro callitrix
30	169.5	8.6	147	4	09OH58	09oh58 homo sapien
31	167.5	8.5	118	6	09MZB1	09mzb1 ovls arles
32	167.5	8.5	146	11	09OXG6	09oxg6 rattus norv
33	166.5	8.5	150	11	054881	054881 xenopus lae
34	163.5	8.3	188	13	042571	042571 brachydantio
35	163.5	8.3	188	13	073682	073682 brachydantio
36	163	8.3	132	14	09YMF3	09ymf3 orf virus.
37	161	8.2	149	6	09XS47	09xs47 bos taurus
38	161	8.2	210	6	029613	029613 felis silve
39	161	8.2	301	5	09VWP6	09vwp6 drosophila
40	160.5	8.2	149	4	09Y6S8	09y6s8 homo sapien
41	158	8.0	158	11	063434	063434 rattus norv
42	156.5	8.0	141	11	070123	070123 mus musculu
43	155.5	7.9	110	11	088911	088911 rattus norv
44	148.5	7.6	123	6	09N1S1	09n1s1 capreolus c
45	147.5	7.5	207	11	064290	064290 mus musculu

ALIGNMENTS

RESULT 1
ID 043915 PRELIMINARY: PRT: 354 AA.
AC 043915;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE GROWTH FACTOR FIGF.
GN FIGF OR VEGF-D.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP MEDLINE-98140120; PubMed-9479493;
RX Rocchigiani M., Testi M., Luddi A., Orlandini M., Franco B.,
RA Rossi E., Ballabio A., Zuffardi O., Oliviero S.;
RT "Human FIGF: cloning, gene structure, and mapping to chromosome Xp22.1
between the FIGA and the GPR genes.";
RL Genomics 47:207-216(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RX MEDLINE-97349118; PubMed-9205122;
RA Yamada Y., Nezu J., Shimane M., Hirata Y.;
RT "Molecular cloning of a novel vascular endothelial growth factor,
VEGF-D.";
RL Genomics 42:483-488(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-98118549; PubMed-9435229;
RA Achen M.G., Jeltsch M., Kuk E., Maekinen T., Vitai A., Wilks A.F.,
ALITALO K., Stacker S.A.;
RT "Vascular endothelial growth factor D (VEGF-D) is a ligand for the
lysoine kinases VEGF receptor 2 (Flk1) and VEGF receptor 3 (Flt4).";
RL Proc. Natl. Acad. Sci. U.S.A. 95:548-553(1998).
DR EMBL: Y12865; CAA73371.1; JOINED.
DR EMBL: Y12865; CAA73371.1; JOINED.
DR EMBL: Y12866; CAA73371.1; JOINED.
DR EMBL: Y12867; CAA73371.1; JOINED.
DR EMBL: Y12868; CAA73371.1; JOINED.


```

Query Match      77.3%; Score 1517.5; DB 11; Length 326;
Best Local Similarity 85.2%; Pred. No. 6.7e-134;
Matches 270; Conservative 20; Mismatches 22; Indels 5; Gaps 1;

QY 1 MYEWMVAVFMALYVOLYOGSSNENHPYK-----RSSOSTLSEEQOIRAAASLEELR 55
    1 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 1 MYGMNAVNTLMSYLYVVOGFSLEHRAVADVLSRSSRVSLESEQOIRAAATLEELQ 60
    1 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 56 ITHSEDMKLMRCRLRLKSFSTMSDRSASRSTRFAATFYDIETLKYIDEEMORTQSPRE 115
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 61 VAHSEWMKLMRCRLRLKSLANVDSRSTRSRTFAATFYDIETLKYIDEEMORTQSPRE 120
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 116 TCVEVASELCKSTNTFFKPCVAVFRCGGCCNEESLICAMTSTSYSKOLFELSVPLTSY 175
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 121 TCVEVASELCKSTNTFFKPCVAVFRCGGCCNEESVCMNTSTSYSKOLFELSVPLTSY 180
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 176 PELVPKVVANHTGCKCLPTAPRHPYSIIRRSIQIPEDRCSHKKLCIPDMLDMSKCKC 235
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 181 PELVPKVVANHTGCKCLPTAPRHPYSIIRRSIQIPEDRCSHKKLCIPDMLDMSKCKC 240
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 236 VIOENPDLGTEHSHLQEPALCGPHMFEDEDRCEVCCTPKPKDLIQHKNCSCECKE 295
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 241 VIOENPDLGTEHSHLQEPALCGPHMFEDEDRCEVCCTPKPKDLIQHKNCSCECKE 300
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 296 SLETCCKOKHKLPHDPC 312
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 301 SLETCCKOKHKLPHDPC 317
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
057352 PRELIMINARY; PRT; 418 AA.
AC 057352;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR.
GN VEGF-C.
OS Ccturnix ccturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Ccturnix.
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98167900; PubMed=9435294;
RA Eichmann A., Corbel C., Jaffredo T., Breant V., Joukov V., Kumar V.,
RA Allitalo K., Le Douarin N.M.;
RT "Avian VEGF-C: cloning, embryonic expression pattern and stimulation
RT of the differentiation of VEGFR2-expressing endothelial cell
RT precursors."
RL Development 125:743-752(1998).
DR EMBL; Y15837; CAA57599.1; -.
DR HSSP; P15692; IVP.
DR InterPro: IPR000072; -.
DR pfam: PF00341; PDGF; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR PRODOM; PD001629; -. 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS50278; PDGF_2; 1.
DR SMART; SM00141; PDGF; 1.
KW Signal.
FT SIGNAL.
FT CHAIN 1 31 POTENTIAL.
SO SEQUENCE 418 AA; 46839 MW; 099BFC79151BP2B CRC64;

```

```

Query Match      35.3%; Score 692; DB 13; Length 418;
Best Local Similarity 39.4%; Pred. No. 1.1e-56;
Matches 140; Conservative 52; Mismatches 101; Indels 62; Gaps 9;

QY 41 EQOIRAAASLEELRLTHSEDMKLMRCRL-----KSFTMSDRSASHSTRFAATFY 94
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 56 EQOLRSVSSVDELMTVLYPEYWMKFCQLRKGGWQHNRHSSDTRDD--SLKFAAHY 113
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 95 DIETLKYIDEEMORTQSPRETCVEVASELCKSTNTFFKPCVAVFRCGGCCNEESLICM 154
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 114 NAILKSIDEMRKQTQCMREVCYDVQKEGATNTFFKPCYSIRGGCCSBLQCM 173
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 155 NTSTSYSKOLFELSVPLTSVPELVKVVANHTGCKCLPTAP--RHPYSIIRRSIQIPRE 212
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 174 NISTNYISKLFELIYPLSHGPKPVYSFANHTSCROMSKLDVYQVHSHIRSLP-ATQ 232
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 213 DRCSHKKLCIPDMLDMSKCKCVLOEE--NLAGED----- 248
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 233 TQCHVANKTCQPKNVVNNQICRLAQHDFGSSHLSDSTSEGFHICGPKLEDEETQC 292
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 249 -----HSHLQ-----PALCGPHMFEDEDRCEVCCTPKPKDLQ 283
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 293 VCKGVRPISCGPHKELDRASCOCCKNKLPPSSCGPKNFEDEKQCYCKKTCQPKHPL 352
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 284 HPRNCSFECKESLETCCKOKHKLPHDPCG-EDRCPHTRPCASGKTACAKHCR 337
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 353 NPAKCIQ-ECTESPNCFLKGRFHQTCSCYRPTVTRKCDAGFLLAEEYCR 406
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
09XS50 PRELIMINARY; PRT; 420 AA.
AC 09XS50;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEART;
RA Liu X., Yonekura H., Yamagishi S., Yamamoto Y., Yamamoto H.;
RT "Structure and expression of bovine VEGF family."
RL Submitted (MAY-1997) to the EMBL/Genbank/DBD databases.
DR EMBL; AB004275; BAAT7687.1; -.
DR HSSP; P15692; IVP.
DR InterPro: IPR000072; -.
DR pfam: PF00341; PDGF; 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS50278; PDGF_2; 1.
DR SMART; SM00141; PDGF; 1.
KW Signal.
FT SIGNAL.
FT CHAIN 1 20 POTENTIAL.
SO SEQUENCE 420 AA; 46681 MW; 58BA84317A3C8E2D CRC64;

```

[illegible]

Query Match	14.6%	Score 287;	DB 11;	Length 126;
Best Local Similarity	44.2%	Pred. No. 2.2e-19;		
Matches 57;	Conservative 24;	Mismatches 40;	Indels 8;	Gaps 4;
DR EMBL: AF010302; AAB63248.1; -				
DR HSSP: P15692; 2YPE				
DR InterPro: IPR000072;				
DR Pfam: PF00341; PDGF_1				
DR PROSITE: PS0278; PDGF_2; 1				
DR SMART: SM00141; PDGF_1				
FT NON_TER 1				
FT NON_TER 126				
SO SEQUENCE 126 AA; 13977 MW; 8f365AFBC4E037B0 CRC64;				
Query Match	14.6%	Score 287;	DB 11;	Length 126;
Best Local Similarity	44.2%	Pred. No. 2.2e-19;		
Matches 57;	Conservative 24;	Mismatches 40;	Indels 8;	Gaps 4;
DR EMBL: AF010302; AAB63248.1; -				
DR HSSP: P15692; 2YPE				
DR InterPro: IPR000072;				
DR Pfam: PF00341; PDGF_1				
DR PROSITE: PS0278; PDGF_2; 1				
DR SMART: SM00141; PDGF_1				
FT NON_TER 1				
FT NON_TER 126				
SO SEQUENCE 126 AA; 13977 MW; 8f365AFBC4E037B0 CRC64;				
Query Match	10.4%	Score 204;	DB 4;	Length 254;
Best Local Similarity	23.6%	Pred. No. 2.5e-11;		
Matches 70;	Conservative 31;	Mismatches 106;	Indels 90;	Gaps 11;
DR RAASLELLLRTHSEDKIMRCRLKLSFTSMDSRA-----SRSTRFAATFYDI 96				
11:::11				
18 RASETWNPL-----SVVHWSLALLILYLIHAKWSQAAPMAEGGGGNH----- 60				

```

QY 97 ETLKVIDEEMORTQCSPRETCVEVASELGSKTNTFFKPCVNVFRCGGCCNEESLICMNT 156
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 61 EYVKFMD-VYQSGYFCPIETLVIDIFOEYDELEIFKPCVPLMRGGCCNDGSLCVP 119
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 157 STSYISKQLEISVPLTSP-----ELVPKAVANHGGCKCLPTAPRHPYSIIRRSIOPEE 212
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 120 EESNITMQLIRIK-----PHOGHIGEMSFLQHNKCECR-----PKK 156
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 213 DRCSHKKICLPIDMLDSNK-----CKCVIQEENPLAGTEDHSHLOEPALC 258
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 157 DRRQREKSVYRGKGQKRRKRSKYSVYVGARCCCLPMS-----LPGPHPC 206
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 259 GPHMDEDRCEVCCTPCPKDLOHPKNSCFECKESLETCCQKHLPDPSCSE 315
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 207 GP-----CSERRKHLFVODPOTCKC-SCKNITDSRCAROLELNERTCRD 250
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 9
Q9GKRO PRELIMINARY: PRT; 190 AA.
AC 09GKRO:
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR 165.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RA SEQUENCE FROM N.A.
RA Miura N., Misumi K., Kawahara K., Nakashima M., Fukumitsu S.,
   Kawabata H., Uto N., Oka T., Maruyama I., Sakamoto H.;
   "Cloning of cDNA and High-Level Expression of Equine Vascular
   Endothelial Growth Factor (VEGF).";
   Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
RL EMBL; AB053350; BAB20890.1;
SQ SEQUENCE 190 AA; 22312 MW; 87E9E1614395F87 CRC64;

Query Match 10.3%; Score 203; DB 6; Length 190;
Best Local Similarity 23.3%; Pred. No. 2.3e-11;
Matches 60; Conservative 29; Mismatches 87; Indels 82; Gaps 9;

QY 62 WKLMRCRLFKSPTSMDSRA-----SHRSTRFAATFYDIETLKVIDEEMORTQCSPRETC 117
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 7 WYHWSLALLLHHAHAKWSOAAPMAEGEHKTHEYVKFMDV-----YQRSICPIETL 57
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 118 VEVASELGSKTNTFFKPCVNVFRCGGCCNEESLICMNTSTSYISKQLEISVPLTSP 177
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 58 VDFQYEPDEIEYIFKPCVPLMRGGCCNDGSLCVPAPAEFNITMQLIRIK-PHQS-QH 115
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 178 LVPKAVANHGGCKCLPTAPRHPYSIIRRSIOPEEDRCHSKSKKCLPIDMLDSNKCVL 237
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 116 IGEMSFLOHSHKCECR-----PKKDKA----- 136
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 238 OEENPLAGTEDHSHLOEPALCGPHMFDDEDRCEVCCTPCPKDLOHPKNSCFECKESL 297
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 137 ROENP-----CGP-----CSERRKHLFVODPOTCKC-SCKNITD 168
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 298 ETCCQKHLPDPSCSE 315
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 169 SRCKAROLELNERTCRD 186
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 10
ID 077643 PRELIMINARY: PRT; 190 AA.
AC 077643:
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR.

```

```

GN VEGF.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RA SEQUENCE FROM N.A.
RA STRAIN-COLUMBIA-RAMBOULIET;
RC Cheung C.Y., Grace R.A.;
RT "Ovine vascular endothelial growth factor: Nucleotide sequence and
   expression in fetal tissues.";
RL Growth Factors 0:0-0(1998).
DR EMBL; AF071015; AAC23608.1;
DR HSSP; P15692; 1YGH.
DR InterPro; IPR000072;
DR Pfam; PF00341; PDGF_1;
DR ProDom; PD001629;
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS0278; PDGF_2; 1.
DR SMART; SM00141; PDGF; 1.
SQ SEQUENCE 190 AA; 22342 MW; 0D5E3B3E5C53E739 CRC64;

Query Match 10.2%; Score 201; DB 6; Length 190;
Best Local Similarity 25.4%; Pred. No. 3.6e-11;
Matches 57; Conservative 21; Mismatches 66; Indels 80; Gaps 8;

QY 97 ETLKVIDEEMORTQCSPRETCVEVASELGSKTNTFFKPCVNVFRCGGCCNEESLICMNT 156
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 38 EYVKFMD-VYQSGYFCPIETLVIDIFOEYDELEIFKPCVPLMRGGCCNDGSLCVP 96
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 157 STSYISKQLEISVPLTSP-----ELVPKAVANHGGCKCLPTAPRHPYSIIRRSIOPEEDRCS 216
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 97 EEFNITMQLIRIK-----PHOSQ-----IGEMSFLO 123
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 217 HSKKICLPIDMLDSNKCKC-----VLQENPLAGTEDHSHLOEPALCGPHMFDDEDRCE 271
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 124 H-----NKCECRPKKDKARQENP-----CGP----- 144
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 272 VCKTPCPKDLOHPKNSCFECKESLETCCQKHLPDPSCSE 315
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 145 -CSERRKHLFVODPOTCKC-SCKNITDSRCAROLELNERTCRD 186
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 11
ID 060720 PRELIMINARY: PRT; 209 AA.
AC 060720:
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE VEGF183 PROTEIN PRECURSOR (VASCULAR ENDOTHELIAL GROWTH FACTOR 183)
   (DJ261G23.6.6) (VASCULAR ENDOTHELIAL GROWTH FACTOR).
GN VEGF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RA SEQUENCE FROM N.A.
RA TISSUE-KIDNEY;
RX MEDLINE=99096474; PubMed=9878851;
RA Lei J., Jiang A., Pel D.;
RT "Identification and characterization of a new splicing variant of
   vascular endothelial growth factor: VEGF183.";
RL Biochim. Biophys. Acta, Gene Struct. Expr. 1443:400-406(1998).
RN [2]
RP SEQUENCE OF 114-209 FROM N.A.
RA TISSUE-RETINA;
RC Dingjing L., Rogue R.S.;
RL Submitted (MAY-1998) to the EMBL/Genbank/DBJ databases.
RN [3]

```


XX	AAW49036		
XX	AAW49036	standard; Protein: 354 AA.	
XX	AC		
XX	AAW49036;		
XX	26-OCT-1998	(first entry)	
DE	Human zwegf2 growth factor.		
KM	Human zwegf2 growth factor; mitogen; fibroblast;		
KM	venous stasis ulcer; diabetic ulcer; skin wound;		
KM	angiogenic effect; tumour; diabetic retinopathy;		
KM	scleroderma.		
OS	Homo sapiens.		
XX	Key	Location/Qualifiers	
FT	Peptide	1..23	
FT		/note= "signal peptide"	
FT	Peptide	24..108	
FT		/note= "Pro-region"	
FT	Binding-site	109..197	
FT		/note= "Receptor binding domain"	
Region		206..256	
FT		/note= "Cysteine-rich domain"	
FT	Region	257..274	
FT		/note= "Balbiant ring motif"	
FT	Region	275..294	
FT		/note= "Balbiani ring motif"	
FT	Region	295..354	
FT		/note= "Cysteine-rich domain"	
XX			
NN	WO9824811-A2.		

XX 11-JUN-1998.
 PD 20-NOV-1997; 97WO-US20888.
 XX 18-SEP-1997; 97US-0933455.
 PR 06-DEC-1996; 96US-0759657.
 XX (ZYMO) ZYMOGENETICS INC.
 PA Conklin DC, Gilbert T, Hart CE, Nygaard S, Sheppard PO.
 PI WPI: 1998-333256/29.
 DR N-PSDB; AAV32823.
 XX New isolated vascular endothelial growth factor - used to develop
 PT products for treating e.g. wounds, burns, myocardial infarction,
 PT tumours, psoriasis, arthritis, restenosis or organ transplants
 XX Claim 1; Pages 53-54; 77pp; English.
 CC The present sequence represents a human zvegf2 growth factor encoded
 CC by the zvegf2 cDNA which was isolated from a human heart cDNA library.
 CC zvegf2 protein in a dimeric form acts as a mitogen for fibroblasts or
 CC smooth muscle cells. zvegf2 is claimed to be useful for stimulating the
 CC revascularisation of tissue or the re-endothelialisation of vascular
 CC tissue. zvegf2 is particularly claimed to be useful for the treatment
 CC of full-thickness skin wounds, including venous stasis ulcers and
 CC diabetic ulcers. The zvegf2 protein is also claimed to be useful as an
 CC additive in tissue adhesives for promoting revascularisation of the
 CC healing tissue. Antagonists against zvegf2 can be used to block its
 CC mitogenic, chemotactic and angiogenic effects. The antagonists may
 CC therefore be useful for reducing growth of solid tumours by inhibiting
 CC neovascularisation of the developing tumour or by directly blocking
 CC tumour cell growth, in the treatment of diabetic retinopathy, psoriasis,
 CC arthritis, and scleroderma.
 CC Sequence 354 AA:
 SQ
 Query Match 100.0%; Score 468; DB 19; Length 354;
 Best Local Similarity 100.0%; Pred. No. 8.3e-48;
 Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 MYREWVVVVFMILYVOLVOGSSNEHGPVKRSSQSTLERSEOOIRASSLEELLRTITSE 60
 Db 1 myrewvvvvfmillyvolvgssnehgpykrssqstlerseqqiraasleeellrtitsh 60
 OY 61 DWKLMRCRLRLKSFSTMSDRSASHRSTRPAAT 92
 Db 61 dwklwrcrlrlksfstsmdrsasshrstrfaat 92
 RESULT 2
 AAM53241 standard; Protein: 354 AA.
 ID AAM53241;
 AC AAM53241;
 XX 03-AUG-1998 (first entry)
 DT Homo sapiens vascular endothelial growth factor D (VEGF-D).
 DE Homo sapiens vascular endothelial growth factor D (VEGF-D).
 XX vascular endothelial growth factor; VEGF-D; angiogenesis;
 KW modification; acceleration; wound healing; tissue; organ;
 KW transplants; collateral circulation; infarction; arterial stenosis;
 KW coronary artery disease; inhibition; cancer; treatment;
 KW diabetic retinopathy; lung disorders; blood circulation;
 KW gaseous exchange; chronic obstructive airway disease;
 KW intestinal malabsorptive syndrome; biopsy; metastatic risk;
 KW detection; diagnosis; congestive heart failure.
 XX Homo sapiens.
 OS

XX WO9807832-A1.
 PN 26-FEB-1998.
 XX 21-AUG-1997; 97WO-US14696.
 PF 01-JUL-1997; 97US-0051426.
 PR 23-AUG-1996; 96AU-0001825.
 PR 23-AUG-1996; 96US-0023751.
 PR 11-NOV-1996; 96AU-0003554.
 PR 14-NOV-1996; 96US-0031097.
 PR 05-FEB-1997; 97AU-0004954.
 PR 10-FEB-1997; 97US-0038814.
 PR 19-JUN-1997; 97AU-0007435.
 XX (LUDM-) LUDMIG INST CANCER RES.
 PA (UYHE-) UNITV HELSINKI LICENSING LTD.
 PI Achen MG, Allitalo K, Stackner SA, Wilks AF;
 XX WPI: 1998-179057/16.
 DR N-PSDB; AAV20807.
 XX New isolated vascular endothelial growth factor-D - used to develop
 PT products for use in e.g. modifying angiogenesis or treating lung,
 PT heart or intestinal disorders
 XX Claim 16; Pages 60-61; 101pp; English.
 PS The sequence is that of human lung vascular endothelial growth factor
 CC D (VEGF-D). VEGF-D can be used for e.g. acceleration of angiogenesis
 CC in wound healing, tissue or organ transplantation, or to establish
 CC collateral circulation in tissue infarction or arterial stenosis,
 CC such as coronary artery disease, and inhibition of angiogenesis in
 CC the treatment of cancer or of diabetic retinopathy. It can also be
 CC used in the treatment of lung disorders to improve blood circulation
 CC in the lung and/or gaseous exchange between the lungs and the blood
 CC stream or to improve blood circulation to the heart and O2 gas
 CC permeability in cases of cardiac insufficiency, to improve blood
 CC flow and gaseous exchange in chronic obstructive airway disease,
 CC or to treat malabsorptive syndromes in the intestinal tract.
 CC Quantitation of VEGF-D in cancer biopsy specimens may be useful
 CC as an indicator of future metastatic risk. Antagonists can be used
 CC for treating e.g. conditions such as congestive heart failure,
 CC involving accumulations of fluid in the lung resulting from
 CC increases in vascular permeability. The products can also be used
 CC for detection and diagnosis.
 CC Sequence 354 AA:
 SQ
 Query Match 100.0%; Score 468; DB 19; Length 354;
 Best Local Similarity 100.0%; Pred. No. 8.3e-48;
 Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 MYREWVVVVFMILYVOLVOGSSNEHGPVKRSSQSTLERSEOOIRASSLEELLRTITSE 60
 Db 1 myrewvvvvfmillyvolvgssnehgpykrssqstlerseqqiraasleeellrtitsh 60
 OY 61 DWKLMRCRLRLKSFSTMSDRSASHRSTRPAAT 92
 Db 61 dwklwrcrlrlksfstsmdrsasshrstrfaat 92
 RESULT 3
 AAM44293 standard; Protein: 354 AA.
 ID AAM44293;
 AC AAM44293;
 XX 22-JUN-1998 (first entry)
 DT Homo sapiens vascular endothelial growth factor D (VEGF-D).
 DE Homo sapiens vascular endothelial growth factor D (VEGF-D).
 XX vascular endothelial growth factor; VEGF-D; angiogenesis;
 KW modification; acceleration; wound healing; tissue; organ;
 KW transplants; collateral circulation; infarction; arterial stenosis;
 KW coronary artery disease; inhibition; cancer; treatment;
 KW diabetic retinopathy; lung disorders; blood circulation;
 KW gaseous exchange; chronic obstructive airway disease;
 KW intestinal malabsorptive syndrome; biopsy; metastatic risk;
 KW detection; diagnosis; congestive heart failure.
 XX Homo sapiens.
 OS

DE Human vascular endothelial growth factor D.
 XX
 KW Human; vascular endothelial growth factor D; VEGF-D; gene therapy;
 XX inflammation; oedema.
 XX
 OS Homo sapiens.
 PN WO9802543-A1.
 XX
 PD 22-JAN-1998.
 XX
 PF 15-JUL-1997; 97WO-JP02456.
 XX
 PR 15-JUL-1996; 96JP-0185216.
 XX
 PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
 XX
 PI Hirata Y, Nezu J;
 XX
 DR WPI; 1998-110591/10.
 DR N-PSDB; AAV15156.
 XX
 PT VEGF-D protein encoded by DNA - useful for, e.g. gene therapy and
 XX treating oedema
 PS
 CC Claim 1; Page 18-20; 52pp; Japanese.
 XX
 CC The present sequence represents human vascular endothelial growth factor
 CC D (VEGF-D). The VEGF-D protein, compounds and antibodies, which can bind
 CC the protein, may be useful in, e.g. gene therapy and in treatment of
 CC inflammation and oedema. Vectors, containing the VEGF-D DNA, and VEGF-D
 CC DNA sequences may be used for screening for the compounds which bind to
 CC the VEGF-D protein.
 XX
 SQ Sequence 354 AA:
 XX
 Query Match 100.0%; Score 468; DB 19; Length 354;
 Best Local Similarity 100.0%; Pred. No. 8.3e-48;
 Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 MYREVVVVVFMMLYVOLVQSSNEHGPVKRSSOSTLERSEQOIRASSLEELRTTSE 60
 DB 1 MYREVVVVVFMMLYVOLVQSSNEHGPVKRSSOSTLERSEQOIRASSLEELRTTSE 60
 OY 61 DKMLRCRLRLKSFSTMSDRSASHSTRPAAT 92
 DB 61 DKMLRCRLRLKSFSTMSDRSASHSTRPAAT 92
 DB 61 dkwkwrclrlrlksfstsmdrsashstrfpaat 92
 RESULT 4
 AAB10649
 ID AAB10649 standard; Protein; 354 AA.
 XX
 AC AAB10649;
 XX
 DF 19-JAN-2001 (first entry)
 XX
 DE Human VEGD protein.
 XX
 KW VEGF-X; vascular endothelial growth factor; human; vulnery; cytostatic;
 KW anti-rheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
 KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
 KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
 KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
 KW venous sore; diabetic ulcer; burns; skin graft growth; VEGD.
 XX
 OS Homo sapiens.
 XX
 PN WO200037641-A2.
 XX
 PD 29-JUN-2000.
 XX

PF 21-DEC-1999; 99WO-US30503.
 XX
 PR 22-DEC-1998; 98GB-0028377.
 XX
 PR 18-MAR-1999; 99US-0124967.
 XX
 PR 08-NOV-1999; 99US-0164131.
 XX
 PA (JANC) JANSSEN PHARM NV.
 XX
 PI Gordon RD, Sprengel JJ, Von JR, Dijkmans JJH, Gosiewska A;
 PI Dhanaraj SN, Xu J;
 XX
 DR WPI; 2000-442669/38.
 XX
 PT New vascular endothelial growth factor protein, useful for treating or
 PT preventing diseases associated with inappropriate angiogenesis activity
 PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
 XX
 XX Disclosure; Fig 11; 127pp; English.
 XX
 CC This invention describes a novel vascular endothelial growth factor-X
 CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
 CC vulnery, cytostatic, anti-rheumatic, antiarthritic, antipsoriatic and
 CC anti-diabetic activity and acts as an angiogenesis and vascularization
 CC regulator. An antisense molecule of the invention is useful for treating
 CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
 CC retinopathy by inhibiting angiogenic activity or inappropriate
 CC vascularization including formation and proliferation of new blood
 CC vessels, growth and development of tissues, tissue regeneration and organ
 CC and tissue repair in a subject. The products of the invention are useful
 CC for preparing medicaments for treating wounds such as dermal ulcers,
 CC pressure sores, venous sores, diabetic ulcers and burns and to promote
 CC skin graft growth, tissue repair, proliferation of new blood vessels,
 CC tissue regeneration and organ repair by promoting angiogenic activity or
 CC vascularization. This sequence represents the human VEGD protein used
 CC to illustrate the method of the invention.
 XX
 SQ Sequence 354 AA:
 XX
 Query Match 100.0%; Score 468; DB 21; Length 354;
 Best Local Similarity 100.0%; Pred. No. 8.3e-48;
 Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 MYREVVVVVFMMLYVOLVQSSNEHGPVKRSSOSTLERSEQOIRASSLEELRTTSE 60
 DB 1 MYREVVVVVFMMLYVOLVQSSNEHGPVKRSSOSTLERSEQOIRASSLEELRTTSE 60
 OY 61 DKMLRCRLRLKSFSTMSDRSASHSTRPAAT 92
 DB 61 DKMLRCRLRLKSFSTMSDRSASHSTRPAAT 92
 DB 61 dkwkwrclrlrlksfstsmdrsashstrfpaat 92
 RESULT 5
 AAB29049
 ID AAB29049 standard; Protein; 354 AA.
 XX
 AC AAB29049;
 XX
 DF 31-JAN-2001 (first entry)
 XX
 DE Human VEGF-D protein sequence.
 XX
 KW Human; Flt4; fms-like tyrosine kinase 4; lymphoedema;
 KW vascular endothelial growth factor receptor 3; VEGFR-3;
 KW Milroy-Nonne syndrome; lymphoedema praecox; VEGF-D;
 KW vascular endothelial growth factor D.
 XX
 OS Homo sapiens.
 XX
 PN WO200058511-A1.
 XX
 PD 05-OCT-2000.
 XX

PF 26-MAR-1999; 99WO-US06133.
XX
PR 26-MAR-1999; 99WO-US06133.
XX
PA (LUDM-) LUDMIG INST CANCER RES.
PA (UYHE-) UNIV HELSINKI LICENSING LTD OY.
PA (UYPI-) UNIV PITTSBURGH.
XX
PI Ferrell RE, Altalo K, Finegold DN, Karkkainen M;
XX
DR WPI; 2000-679298/66.
DR N-PSDB; AAC62407.
XX
PT Screening a human subject for increased risk of developing a lymphatic
PT disorder, comprises assaying a nucleic acid to determine a mutation
PT altering the sequence of a vascular endothelial growth factor
PT receptor-3 -
XX
XX
XX Disclosure; Page 64-65; 76pp; English.
XX
CC The present sequence is the protein sequence of the human vascular
CC endothelial growth factor D (VEGF-D). It was used to demonstrate the
CC methods of the invention, which involve the screening of individuals to
CC determine which vascular endothelial growth factor receptor 3 (VEGFR-3,
CC also known as Flt4 or fms-like tyrosine kinase 4) alleles they possess
CC and thus their likelihood of developing hereditary lymphoedema.
CC Conditions associated with lymphoedema include Milroy-Nonne syndrome,
CC which is early onset lymphoedema and lymphoedema praecox, which is late
CC onset.
XX
SQ Sequence 354 AA;
XX
XX
Query Match 100.0%; Score 468; DB 21; Length 354;
Best Local Similarity 100.0%; Pred. No. 8.3e-48;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 MYREVVVVVFMALYYVOLVOGSSNEHGPVKRSSQSTLERSEDOIRAASSLEELLRTITHE 60
Db 1 myrewvvvvfmallyvqlvgssnehgpkvkrssqstlerseqgiraassleellritthe 60
XX
OY 61 DKMLRCRLRLKSFSTMSDSRASHSTRPAAT 92
Db 61 dwklwrcrlrlksfstsmdsrsashstrfaat 92
XX
RESULT 6
AA70750
ID AAY70750 standard; Protein; 354 AA.
AC AAY70750;
XX
DT 17-AUG-2000 (first entry)
XX
DE Human prepro-vascular endothelial growth factor D.
XX
XX Human; receptor tyrosine kinase; RTK; Flt4; fms-like tyrosine kinase 4;
KW VEGFR-3; vascular endothelial growth factor receptor-3; chromosome 5q35;
KW cytostatic; tumour imaging; anti-tumour therapy; treatment; diagnosis;
KW neoplastic disease; lymphoma; carcinoma; breast; squamous cell; melanoma;
KW sarcoma; malignancy; VEGF-D; vascular endothelial growth factor D.
XX
XX Homo sapiens.
OS
XX
PN WO200021560-A1.
XX
PD 20-APR-2000.
XX
XX
PF 08-OCT-1999; 99WO-US23525.
XX
PR 09-OCT-1998; 98US-0169079.
XX
PA (LUDM-) LUDMIG INST CANCER RES.

PA (UYHE-) UNIV HELSINKI LICENSING LTD OY.
XX
XX Alltalo K, Kaipainen A, Vaillola R, Jussila L;
XX
XX WPI; 2000-317650/27.
XX
DR
XX
XX Treating neoplastic diseases such as lymphoma, carcinomas, melanomas
PT and sarcomas, involves administering a compound capable of inhibiting
PT binding of ligand proteins to fms-like tyrosine kinase-4 receptor
XX
XX Example 15-17; Page 142-143; 148pp; English.
XX
XX
XX The patent discloses a method to treat neoplastic disease characterised
XX by expression of fms-like tyrosine kinase 4 (Flt4) receptor (also
XX referred as vascular endothelial growth factor receptor-3, VEGFR-3) in
XX endothelial cells of blood vessels adjacent to malignant neoplasm. The
XX method involves administering a compound that inhibits binding of a
XX ligand to Flt4 thereby inhibiting Flt4 mediated proliferation of vascular
XX endothelial cells. The compound is useful for treating neoplastic disease
XX such as breast carcinomas, squamous cell carcinomas, lymphomas, melanomas
XX and sarcomas. Flt4 receptor tyrosine kinase binding compounds can be used
XX for manufacturing medicament useful for diagnostic screening, imaging and
XX treatment of malignancies characterised by Flt4-expressing blood cells.
XX The Flt4 gene maps to chromosomal region 5q35 and is expressed as 5.8 kb
XX and 4.5 kb mRNAs which differ in their 3' sequences and are
XX differentially expressed in HEL and DAMI cell lines. Flt4
XX belongs to a subfamily of class III receptor tyrosine kinases (RTKs).
XX It is used as a target for tumour imaging and anti-tumour therapy.
XX The present sequence is a human prepro-vascular endothelial growth
XX factor D (VEGF-D), a specific example of Flt4 binding compound. A
XX recombinantly matured VEGF-D lacking residues 1-92 and 202-354
XX retains the ability to activate VEGFR-2 and VEGFR-3 receptors and
XX associate as non-covalently linked dimers.
XX
SQ Sequence 354 AA;
XX
XX
Query Match 100.0%; Score 468; DB 21; Length 354;
Best Local Similarity 100.0%; Pred. No. 8.3e-48;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 MYREVVVVVFMALYYVOLVOGSSNEHGPVKRSSQSTLERSEDOIRAASSLEELLRTITHE 60
Db 1 myrewvvvvfmallyvqlvgssnehgpkvkrssqstlerseqgiraassleellritthe 60
XX
OY 61 DKMLRCRLRLKSFSTMSDSRASHSTRPAAT 92
Db 61 dwklwrcrlrlksfstsmdsrsashstrfaat 92
XX
RESULT 7
AA70983
ID AAY70983 standard; Protein; 354 AA.
AC AAY70983;
XX
DT 09-AUG-2000 (first entry)
XX
DE Human vascular endothelial growth factor (VEGF)-D protein.
XX
XX Vascular endothelial growth factor-D; VEGF; human; re-endothelialisation;
KW vascular endothelial growth factor receptor; VEGFR; vascular trauma;
KW blood vessel; cardiovascular surgery; anti-restenosis agent; prevention;
KW restenosis; stenosis; percutaneous transluminal coronary angioplasty.
XX
XX Homo sapiens.
OS
XX
XX
XX Key Location/Qualifiers
FH 1..21
FT Peptide /label= Signal-peptide
FT /note= "Cleavage results in partially-processed VEGF-D
FT protein"
FT 22..92
FT Peptide

```
FT FT /label= Amino_terminal_peptide
FT FT /note= "Cleavage results in fully-processed mature
FT FT VEGF-D protein"
FT FT 93..201
FT FT /label= Recombinantly-matured_VEGF-D_protein
FT FT /note= "Processed vascular epithelial growth factor-D"
FT FT 202..354
FT FT /label= Carboxy-terminal_peptide
FT FT /note= "Cleavage results in partially-processed VEGF-D
FT FT protein"
FT FT
XX XX WO200024412-A2.
XX XX
XX XX 04-MAY-2000.
XX XX
XX XX 26-OCT-1999; 99WO-US24054.
XX XX
XX XX 26-OCT-1998; 98US-0105587.
XX XX
XX XX (LUDM-) LUDWIG INST CANCER RES.
XX XX (UYHE-) UNIV HELSINKI LICENSING LTD OY.
XX XX (YLAH/) YLA-HERTUUALA S.
XX XX
XX XX Yla-heretuuala S, Aitalo K, Hiltunen MO, Jeltsch MM, Achen MG;
XX XX
XX XX WPI; 2000-350584/30.
XX XX
XX XX N-PSDB; AAD00340.
XX XX
XX XX Preventing stenosis and restenosis in mammals using vascular
XX XX endothelial growth factor proteins or the nucleic acids encoding them -
XX XX
XX XX Disclosure: Page 53-55; 61pp; English.
XX XX
XX XX The present amino acid sequence is the complete human prepro-vascular
XX XX endothelial growth factor (VEGF)-D. VEGF-D has the ability to stimulate
XX XX re-endothelialisation of an injured blood vessel, without significant
XX XX stimulation of smooth muscle cell proliferation. It can bind to and
XX XX stimulate VEGFR-2 (vascular endothelial growth factor receptor) and/or
XX XX VEGFR-3 phosphorylation in cells that express such receptors. An
XX XX anti-restenosis agent comprising either a VEGF-D gene or protein is
XX XX used in a method to reduce or prevent restenosis and stenosis of a blood
XX XX vessel following vascular trauma e.g., cardiovascular surgery and
XX XX percutaneous transluminal coronary angioplasty.
XX XX
XX XX Sequence 354 AA:
XX XX
XX XX
XX XX Query Match 100.0%; Score 468; DB 21; Length 354;
XX XX Best Local Similarity 100.0%; Pred. No. 8.3e-48;
XX XX Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX XX
XX XX QY 1 MYREVVVVVFMMLYVOLVQSSNEHGPVKRSSOSTLERSEDOIRASSLEELLRTTSHSE 60
XX XX |||||||
XX XX Db 1 myrevvvvvfmmllyvqlvgssnehgpkvkrssgstlerseqgiraasleellrttshse 60
XX XX
XX XX QY 61 DKLMRCRLRLKSFMSDRSASHSRSTRPAAT 92
XX XX |||||||
XX XX Db 61 dkwlmrcrlrlksfmsdrsashstrfpaat 92
XX XX
XX XX
XX XX RESULT 8
XX XX AAY97573
XX XX ID AAY97573 standard; Protein; 354 AA.
XX XX
XX XX AAY97573;
XX XX
XX XX 05-APR-2001 (first entry)
XX XX
XX XX Human VEGF-D1 protein sequence.
XX XX
XX XX Human; angiogenic protein; wound healing; vascular tissue repair;
XX XX peripheral arterial disease; critical limb ischaemia; coronary disease;
XX XX angiogenesis; tumour; inflammation; diabetic retinopathy; psoriasis;
XX XX
```

```
KW KW Rheumatoid arthritis; autoimmune disease; allergy; cancer; therapy;
KW KW infectious disease; neurodegeneration;
KW KW vascular endothelial growth factor-D1; VEGF-D1.
XX XX
XX XX Homo sapiens.
XX XX
XX XX WO200075163-A1.
XX XX
XX XX 14-DEC-2000.
XX XX
XX XX 01-JUN-2000; 2000WO-US14925.
XX XX
XX XX 03-JUN-1999; 99US-0137796.
XX XX
XX XX (HUMA-) HUMAN GENOME SCI INC.
XX XX
XX XX Rosen CA, Ruben SM, Hu J, Cao L;
XX XX
XX XX WPI; 2001-071057/08.
XX XX
XX XX N-PSDB; AAA91007.
XX XX
XX XX
XX XX Claim 11; Page 228-229; 244pp; English.
XX XX
XX XX This sequence is vascular endothelial growth factor-D1 (VEGF-D1),
XX XX which is an angiogenic protein of the invention. The angiogenic proteins
XX XX and the DNA sequences encoding them, are used to prevent, treat or
XX XX ameliorate disease and to detect diseases, or susceptibility, by
XX XX detecting mutations or the presence or amount of angiogenic protein
XX XX expression. Particularly they are used to stimulate wound healing,
XX XX growth of damaged bone and tissue, and for repair of vascular tissue,
XX XX especially peripheral arterial disease, critical limb ischaemia or
XX XX coronary disease. Antagonists of the sequences are used to inhibit
XX XX angiogenesis in tumours and to treat inflammation (where associated with
XX XX increased vascular permeability), diabetic retinopathy, rheumatoid
XX XX arthritis or psoriasis. Agonists are also useful for stimulating
XX XX (lymph)angiogenesis. The proteins are also used to identify specific
XX XX binding agents (potential therapeutic agents) and to raise antibodies.
XX XX The antibodies are useful as therapeutic (ant)agonists; for detection,
XX XX purification and targeting of proteins for in vivo or in vitro diagnosis
XX XX (including imaging) or for therapy (including when linked to e.g. a label
XX XX or cytotoxin); and for immunotyping of cells, e.g. for detecting minimal
XX XX residual disease or haematopoietic progenitor/stem cells. It is also
XX XX contemplated that the sequences might be useful for treating a very wide
XX XX range of other disorders, e.g. autoimmune diseases; allergy; cancer;
XX XX infectious diseases (viral, bacterial, fungal or parasitic);
XX XX neurodegeneration, also as chemotactic agents or for stimulating
XX XX regeneration of the nervous system etc.
XX XX
XX XX Sequence 354 AA:
XX XX
XX XX
XX XX Query Match 100.0%; Score 468; DB 22; Length 354;
XX XX Best Local Similarity 100.0%; Pred. No. 8.3e-48;
XX XX Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX XX
XX XX QY 1 MYREVVVVVFMMLYVOLVQSSNEHGPVKRSSOSTLERSEDOIRASSLEELLRTTSHSE 60
XX XX |||||||
XX XX Db 1 myrevvvvvfmmllyvqlvgssnehgpkvkrssgstlerseqgiraasleellrttshse 60
XX XX
XX XX QY 61 DKLMRCRLRLKSFMSDRSASHSRSTRPAAT 92
XX XX |||||||
XX XX Db 61 dkwlmrcrlrlksfmsdrsashstrfpaat 92
XX XX
XX XX
XX XX RESULT 9
XX XX AAB37606
XX XX ID AAB37606 standard; Protein; 354 AA.
XX XX
XX XX AAB37606;
XX XX
```

```

XX 27-FEB-2001 (first entry)
DT
XX
DE Human VEGF-D.
XX
KW Human; gene therapy; lymphatic disorder; hereditary lymphoedema; Flt4;
KW vascular endothelial growth factor receptor-3; VEGFR-3; VEGF-C; VEGF-D;
KW fms-like tyrosine kinase 4.
XX
OS Homo sapiens.
XX
PN CA2283470-A1.
XX
PD 26-SEP-2000.
XX
PF 29-SEP-1999; 99CA-2283470.
XX
PR 26-MAR-1999; 99MO-US06133.
XX
PR 16-AUG-1999; 99US-0375248.
XX
PA (UYP1-) UNIV PITTSBURGH.
PA (UYHE-) UNIV HELSINKI LICENSING LTD OY.
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Alltalo K, Ferrell RE, Finegold DN, Karkkainen M;
XX
DR WPI: 2001-007762/02.
XX
DR N-PSDB; AAC68954.
XX
PT Screening a human for an increased risk of developing lymphatic
PT disorder comprising assaying nucleic acid for alterations in the
PT sequences expressing vascular endothelial growth factor receptor-3
XX
PS Disclosure: Pages 66-67; 99pp; English.
XX
CC The present invention relates to a method for screening a human subject
CC for an increased risk of developing a lymphatic disorder e.g. hereditary
CC lymphoedema. The method comprises assaying nucleic acid of a human
CC subject to determine a presence or an absence of a mutation altering the
CC sequence or expression of vascular endothelial growth factor receptor-3
CC (VEGFR-3)/fms-like tyrosine kinase 4 (Flt4) allele (see AAC68952 and
CC AAB7604) and determining an increased risk of developing lymphatic
CC disorder from presence or absence of the mutation. The presence of a
CC mutation altering the encoded amino acid sequence or expression of at
CC least 1 VEGFR-3 allele in the nucleic acid correlates with an increased
CC risk of developing a lymphatic disorder. Treatment for hereditary
CC lymphoedema can be provided through the administration of vascular
CC endothelial growth factor C (VEGF-C) and vascular endothelial growth
CC factor D VEGF-D genes (via gene therapy) and proteins. The present
CC sequence is the protein sequence for VEGF-D.
XX
SQ Sequence 354 AA:

```

Query Match 100.0%; Score 468; DB 22; Length 354;
 Best Local Similarity 100.0%; Pred. No. 8.3e-48;
 Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MYREVVVVVFMMLYVOLVOGSSNEHGPVKRSSOSTLERSEQOIRAAASLEELLRTTISE 60
DB 1 MYREVVVVVFMMLYVOLVOGSSNEHGPVKRSSOSTLERSEQOIRAAASLEELLRTTISE 60
OY 61 DWKLMRCRLRLKSFSTMSDRSASHRSTRPAAT 92
DB 61 dwklwrcrlrlksfstmsdrsrshstrpaat 92

```

RESULT 10
 AAB70685
 ID AAB70685 standard; protein: 354 AA.
 AC AAB70685;
 XX

```

DT 16-MAY-2001 (first entry)
XX
DE Human vascular endothelial growth factor D (VEGF-D) protein.
XX
KW Human; vascular endothelial growth factor D; VEGF-D; angiogenic;
KW angiogenesis; c-fos induced growth factor; Fgf; cardiant; vasotropic;
KW ischemic condition; ischemia; vascular insufficiency;
KW peripheral vascular disease; coronary artery disease;
KW myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200112669-A1.
XX
PD 22-FEB-2001.
XX
PF 16-AUG-2000; 2000WO-IB01244.
XX
PR 16-AUG-1999; 99US-0149300.
XX
PA (UYST-) UNIV SIENA.
XX
PI Oliviero S;
XX
DR WPI: 2001-202857/20.
XX
PT use of recombinant vascular endothelial growth factor or its
PT angiogenically active fragment or mutant, for inducing angiogenesis in
PT vivo or in vitro and for treating myocardial infarction, coronary
PT artery disease -
XX
PS Claim 2: Page 51-52; 55pp; English.
XX
CC The present invention describes a method for inducing angiogenesis in a
CC tissue, or area, in need of angiogenesis, in a mammal. The method
CC comprises administering recombinant vascular endothelial growth factor D
CC (VEGF-D) or its angiogenically active fragment or mutant (I). The
CC present sequence represents the human VEGF-D protein, which can be used
CC in the method of the invention. (I) has cardiant and vasotropic
CC activities, and is an angiogenesis inducer. The method can be used for
CC inducing angiogenesis in a tissue in need of angiogenesis, or in an area
CC in need of angiogenesis in a mammal. (I) is useful for treating various
CC ischaemic conditions manifested by vascular insufficiency such as
CC peripheral vascular disease, coronary artery disease or myocardial
CC infarction.
XX
SQ Sequence 354 AA:

```

Query Match 98.9%; Score 463; DB 22; Length 354;
 Best Local Similarity 98.9%; Pred. No. 3.3e-47;
 Matches 91; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 MYREVVVVVFMMLYVOLVOGSSNEHGPVKRSSOSTLERSEQOIRAAASLEELLRTTISE 60
DB 1 MYREVVVVVFMMLYVOLVOGSSNEHGPVKRSSOSTLERSEQOIRAAASLEELLRTTISE 60
OY 61 DWKLMRCRLRLKSFSTMSDRSASHRSTRPAAT 92
DB 61 dwklwrcrlrlksfstmsdrsrshstrpaat 92

```

RESULT 11
 AAM14994
 ID AAM14994 standard; Protein: 620 AA.
 AC AAM14994;
 XX
 AC AAM14994;
 XX
 DT 05-JUL-1997 (first entry)
 XX
 DE Human c-Fos induced growth factor (clone HF175 ORF2 product).
 XX
 KW c-Fos induced growth factor; FGF; Fos regulated gene;

KW proto-oncogene; lung disorder; cancer; tumour; therapy;
 XX antibody; transgenic animal.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Misc-difference 16 "residue 16 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT Misc-difference 26 "residue 26 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT Misc-difference 29 "residue 29 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT Misc-difference 47 "residue 47 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT Misc-difference 71 "residue 71 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT Misc-difference 72 "residue 72 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT Misc-difference 76 "residue translated from ORF2 of Hf175
 is Ile"
 FT Misc-difference 136 "residue translated from ORF2 of Hf175
 is Ile"
 FT Misc-difference 220 "residue translated from ORF2 of Hf175
 is Phe"
 FT Misc-difference 341 "residue translated from ORF2 of Hf175
 is His"
 FT Misc-difference 344 "residue translated from ORF2 of Hf175
 is Phe"
 FT Misc-difference 377 "residue translated from ORF2 of Hf175
 is Leu"
 FT Misc-difference 435 "residue 435 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT Misc-difference 486 "residue 486 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT Misc-difference 497 "residue 497 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT Misc-difference 518 "residue 518 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT Misc-difference 541 "residue 541 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT Misc-difference 553 "residue 553 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT Misc-difference 557 "residue 557 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT Misc-difference 562 "residue 562 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT Misc-difference 579 "residue 579 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT Misc-difference 592 "residue 592 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT Misc-difference 593 "residue 593 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"

FT FT "residue 593 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT Misc-difference 597 "residue 597 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT Misc-difference 605 "residue 605 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT Misc-difference 608 "residue 608 corresponds to an in-frame
 stop codon in reading frame 2 of Hf175"
 FT WO9712972-A2.
 XX PN
 XX PD
 XX 10-APR-1997.
 XX PF 30-SEP-1996; 96WO-1B01113.
 XX PR 13-JUN-1996; 96GB-0012368.
 XX PR 29-SEP-1995; 95GB-0019928.
 XX PA (UYSI-) UNIV SIENA.
 XX PI Oliviero S;
 XX DR WPI: 1997-226216/20.
 XX DR N-PSDB: AAT62961.
 XX PS Claim 3; Fig 2; 64pp; English.
 CC 3 Polypeptide sequences (AAW1493-95) are the respective translated
 CC sequences of reading frames 1, 2 and 3 of clone Hf175 (AAT62961), the
 CC human homologue of murine clone F0401 (AAT62960), which codes for a
 CC novel c-Fos induced growth factor (FIGF) (see also AAW14992).
 CC Examination of the 3 polypeptides indicates that reading frame 2
 CC has the longest region free of nonsense codons. FIGF is a c-Fos-
 CC dependent autocrine growth factor able to induce cell division
 CC entry and, when over-expressed, a transformed phenotype in
 CC fibroblasts. It could be implicated in tumours and development.
 CC Recombinant FIGF can be produced in transformed host (e.g. CHO)
 CC cells. It can be used to identify its receptors and in an assay
 CC for the identification of agonists and antagonists. Antibodies
 CC raised against FIGF can be used to block the function of the
 CC protein and thereby inhibit or suppress tumour growth. Transgenic
 CC animals expressing FIGF can be generated for use e.g. as models for
 CC research.
 XX SQ Sequence 620 AA:
 Query Match 98.9%; Score 463; DB 18; Length 620;
 Best Local Similarity 98.9%; Pred. No. 6,5e-47;
 Matches 91; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 MYREVVVVVFPMLYVQLVQSSNEHGPVKRSSQSTLERSQOIRAAASLEILLRTTSHSE 60
 DB 81 MYREVVVVVFPMLYVQLVQSSNEHGPVKRSSQSTLERSQOIRAAASLEILLRTTSHSE 140
 OY 61 DKIMRCRLKTSFTSMDSRSASHSTRPAAT 92
 DB 141 DKLWRCRLIKTSFTSMDSRSASHSTRPAAT 172
 RESULT 12
 ID AAW53243 standard: protein; 321 AA.
 XX AAW53243;
 AC AAW53243;
 XX

Dt		03-AUG-1998	(first entry)
DE	Mus musculus vascular endothelial growth factor D2 (VEGF-D2).		
KW	vascular endothelial growth factor; VEGF-D; angiogenesis;		
KM	modification; acceleration; wound healing; tissue; organ;		
KV	transplants; collateral circulation; infarction; arterial stenosis;		
KX	coronary artery disease; inhibition; cancer; treatment;		
KY	diabetic retinopathy; lung disorders; blood circulation;		
LK	gaseous exchange; chronic obstructive airway disease;		
MM	intestinal malabsorptive syndrome; biopsy; metastatic risk;		
NH	detection; diagnosis; congestive heart failure.		
OS	Mus musculus.		
PB			
PD	WO960783Z-AI.		
PE	26-FEB-1998.		
PF	21-AUG-1997; 97MO-US14696.		
PR	01-JUL-1997; 97US-0051426.		
PR	23-AUG-1996; 96AU-0001825.		
PR	23-AUG-1996; 96US-0023751.		
PR	11-NOV-1996; 96AU-0003554.		
PR	14-NOV-1996; 96US-0031097.		
PR	05-FEB-1997; 97AU-0004954.		
PR	10-FEB-1997; 97US-0038814.		
PS	19-JUN-1997; 97AU-0007435.		
PA	(LUDMW) LUDMWIG INST CANCER RES.		
PA	(UYHE-) UNIV HELSINKI LICENSING LTD.		
PI	Achen MG, Allitalo K, Stacker SA, Wilks AF,		
XX			
DR	WPI; 1998-179057/16.		
DR	N-PSTDB; AAV20809.		
PT	New isolated vascular endothelial growth factor-D - used to develop		
PT	products for use in e.g. modifying angiogenesis or treating lung,		
PP	heart or intestinal disorders		
PS	Claim 16; Pages 64-65; 101pp; English.		
CC	The sequence is that of mouse lung vascular endothelial growth factor		
CC	D2 (VEGF-D2). VEGF-D can be used for e.g. acceleration of angiogenesis		
CC	in wound healing, tissue or organ transplantation, or to establish		
CC	collateral circulation in tissue infarction or arterial stenosis,		
CC	such as coronary artery disease, and inhibition of angiogenesis in		
CC	the treatment of cancer or of diabetic retinopathy. It can also be		
CC	used in the treatment of lung disorders to improve blood circulation		
CC	In the lung and/or gaseous exchange between the lungs and the blood		
CC	stream or to improve blood circulation to the heart and O ₂ gas		
CC	permeability in cases of cardiac insufficiency, to improve blood		
CC	flow and gaseous exchange in chronic obstructive airway disease,		
CC	or to treat malabsorptive syndromes in the intestinal tract.		
CC	Quantitation of VEGF-D in cancer biops specimens may be useful		
CC	as an indicator of future metastatic risk. Antagonists can be used		
CC	for treating e.g. conditions such as congestive heart failure,		
CC	involving accumulations of fluid in the lung resulting from		
CC	increases in vascular permeability. The products can also be used		
CC	for detection and diagnosis.		
XX			
SQ	Sequence 321 AA;		
Ox	Query Match 79.7%; Score 373; DB 19; Length 321; Best Local Similarity 79.3%; Pred. No. 1.6e-36; Matches 73; Conservative 7; Mismatches 12; Indels 0; Gaps 0;		
Bd	1 MYREVVVNVFMLKVOLVGSSNHHGPKRRSSOILRSEGOIRAASTLEELRTTISE 60 : : : : : : 1 myggvgnlnimllmlylvqgfirsnpgvkrrssmlrseseqdltaassleelqlahse 60		

Oy		61 DKKLRRCRLKSFMSDRSASHSTRPAAT 92
Dd		61 dkkLwrcrlklsfmsdrsashtstriaat 92
 RESULT 13 * AAM53242 ID AAM53242 standard; Protein: 358 AA.		
AC		AAM53242;
XX		
DT		03-AUG-1998 (first entry)
XX		
DE		Mus musculus vascular endothelial growth factor D1 (VEGF-D).
XX		
KW		vascular endothelial growth factor; VEGF-D; angiogenesis;
KM		modification; acceleration; wound healing; tissue; organ;
RN		transplants; collateral circulation; infarction; arterial stenosis;
KV		coronary artery disease; inhibition; cancer; treatment;
KW		diabetic retinopathy; lung disorders; blood circulation;
KM		gaseous exchange; chronic obstructive airway disease;
KV		intestinal malabsorptive syndrome; biopsy; metastatic risk;
KX		detection; diagnosis; congestive heart failure.
XX		
OS		Mus musculus.
XX		
PN		WO9807832-A1.
XX		
PD		26-FEB-1998.
XX		
PF		21-AUG-1997; 97MO-USI4696.
XX		
PR		01-JUL-1997; 97US-0051426.
PR		23-AUG-1996; 96AU-0001825.
PR		23-AUG-1996; 96US-0023751.
PR		11-NOV-1996; 96AU-0003554.
PR		14-NOV-1996; 96US-0031097.
PR		05-FEB-1997; 97AU-0004954.
PR		10-FEB-1997; 97US-0038814.
PR		19-JUN-1997; 97AU-0007435.
XX		
PA		(LUDWG-) LUDWIG INST CANCER RES.
PA		(UYHE-) UNIV HELSINKI LICENSING LTD.
XX		
PI		Achen MG, Allitalo K, Stacker SA, Wilks AF;
XX		
DR		WPt: 1998-179057/16.
DR		N-PsDB: AAV20808.
XX		
PT		New isolated vascular endothelial growth factor-D - used to develop
PT		products for use in e.g. modifying angiogenesis or treating lung,
PT		heart or intestinal disorders
XX		
PS		Claim 16; Pages 63-64; 101pp; English.
XX		
CC		The sequence is that of mouse lung vascular endothelial growth factor
CC		D1 (VEGF-D1). VEGF-D1 can be used for e.g. acceleration of angiogenesis
CC		in wound healing, tissue or organ transplantation, or to establish
CC		collateral circulation in tissue infarction or arterial stenosis,
CC		such as coronary artery disease, and inhibition of angiogenesis in
CC		the treatment of cancer or of diabetic retinopathy. It can also be
CC		used in the treatment of lung disorders to improve blood circulation
CC		in the lung and/or gaseous exchange between the lungs and the blood
CC		stream or to improve blood circulation to the heart and O2 gas
CC		permeability in cases of cardiac insufficiency, to improve blood
CC		flow and gaseous exchange in chronic obstructive airway disease,
CC		or to treat malabsorptive syndromes in the intestinal tract.
CC		Quantitation of VEGF-D in cancer biopsy specimens may be useful
CC		as an indicator of future metastatic risk. Antagonists can be used
CC		for treating e.g. conditions such as congestive heart failure,
CC		involving accumulations of fluid in the lung resulting from
CC		increases in vascular permeability. The products can also be used

Wed Oct 17 15:32:53 2001

us-09-427-657-4_copy_1_92.rag

Page 10

Search completed: October 17, 2001, 14:46:03
Job time: 129 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:47:12 ; Search time 62.93 Seconds
(without alignments)
30.102 Million cell updates/sec

Title: US-09-427-657-4_COPY_1_92
Perfect score: 468
Sequence: 1 MYREMYVNVFMMLYVLVQ.....SFTSMDSRSASHRTPRAAT 92

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents_AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/Backfilltest1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	468	100.0	354	4	US-08-915-795-5
2	373	79.7	321	4	US-08-915-795-9
3	360.5	77.0	358	4	US-08-915-795-8
4	309	66.0	325	4	US-08-915-795-3
5	98	20.9	418	4	US-08-795-430-13
6	92.5	19.8	415	4	US-08-795-430-11
7	90	19.2	419	2	US-08-999-811-2
8	90	19.2	419	3	US-09-042-105-2
9	90	19.2	419	3	US-09-042-105-18
10	90	19.2	419	4	US-08-795-430-8
11	90	19.2	419	4	US-08-510-133A-35
12	90	19.2	419	5	PCT-US96-09001-2
13	68	14.5	1271	1	US-08-095-734-2
14	68	14.5	1271	2	US-08-444-623-2
15	68	14.5	1271	3	US-08-471-869-2
16	68	14.5	1271	5	PCT-US94-08267-2
17	59	12.6	148	1	US-08-166-195A-51
18	59	12.6	148	2	US-08-436-772-51
19	59	12.6	148	2	US-08-436-883B-51
20	58	12.4	401	3	US-08-517-802-3
21	57	12.2	976	4	US-09-104-324B-4
22	56.5	12.1	312	2	US-08-518-862C-2
23	56.5	12.1	491	4	US-08-687-916-23
24	56.5	12.1	491	4	US-09-138-614-23
25	56.5	12.1	495	2	US-08-687-916-22
26	56.5	12.1	495	4	US-09-138-614-22
27	55.5	11.9	457	1	US-08-416-478A-8

28	55.5	11.9	457	2	US-08-474-988B-8	Sequence 8, Appl
29	55.5	11.9	457	2	US-08-394-442B-8	Sequence 8, Appl
30	54.5	11.6	206	4	US-09-273-163-7	Sequence 7, Appl
31	54.5	11.6	321	2	US-08-592-126-143	Sequence 143, App
32	54.5	11.6	424	3	US-09-173-581-7	Sequence 7, Appl
33	54.5	11.6	1809	3	US-09-012-515A-12	Sequence 12, Appl
34	54.5	11.6	1809	4	US-08-360-144A-12	Sequence 12, Appl
35	54.5	11.6	2549	5	PCT-US95-06722-12	Sequence 12, Appl
36	54	11.5	243	1	US-07-915-720D-16	Sequence 16, Appl
37	54	11.5	243	3	US-09-025-543-16	Sequence 16, Appl
38	54	11.5	300	1	US-08-218-026-2	Sequence 2, Appl
39	54	11.5	300	2	US-08-653-632-2	Sequence 7, Appl
40	54	11.5	368	1	US-07-915-720D-15	Sequence 15, Appl
41	54	11.5	368	3	US-09-025-543-15	Sequence 15, Appl
42	54	11.5	422	1	US-08-680-726A-68	Sequence 68, Appl
43	54	11.5	422	4	US-09-092-409-68	Sequence 68, Appl
44	54	11.5	487	1	US-08-030-644-2	Sequence 2, Appl
45	54	11.5	487	5	PCT-US93-04754-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-915-795-5
; Sequence 5, Application US/08915795
; Patent No. 6235713
; GENERAL INFORMATION:
; APPLICANT: Marc G. ACHEN
; APPLICANT: Andrew F. WILKS
; APPLICANT: Steven A. STACKER
; APPLICANT: Karl ALITALO
; TITLE OF INVENTION: GROWTH FACTOR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
; STREET: 1200 G Street, NW, Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: United States of America
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,795
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D. 26,269
; REGISTRATION NUMBER: 1064/42983
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; TITEL: N/A
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 354 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: Human Lung
; US-08-915-795-5
Query Match 100.0%; Score 468; DB 4; Length 354;
Best Local Similarity 100.0%; Pred. No. 9.8e-52;

Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYREVVVVVFMALYVOLQGSSNHEGPKVRSOSQTLERSEOOIRASSLEELRITTHSE 60

DB 1 MYREVVVVVFMALYVOLQGSSNHEGPKVRSOSQTLERSEOOIRASSLEELRITTHSE 60

QY 61 DKMLMRCRLRLKSFSTMSDRSASHRSTRFAAT 92

DB 61 DKMLMRCRLRLKSFSTMSDRSASHRSTRFAAT 92

RESULT 2

US-08-915-795-9

Sequence 9, Application US/08915795

Patent No. 6235713

GENERAL INFORMATION:

APPLICANT: Marc G. ACHEN

APPLICANT: Andrew F. WILKS

APPLICANT: Steven A. STACKER

APPLICANT: Kari ALITALO

TITLE OF INVENTION: GROWTH FACTOR

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.

STREET: 1200 G Street, NW, Suite 700

CITY: Washington

STATE: DC

COUNTRY: United States of America

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/915,795

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: EVANS, Joseph D.

REGISTRATION NUMBER: 26,269

REFERENCE/DOCKET NUMBER: 1064/42983

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 628-8800

TELEFAX: (202) 628-8844

TELEX: N/A

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 321 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

TISSUE TYPE: Mouse Lung

US-08-915-795-9

Query Match 79.7%; Score 373; DB 4; Length 321;

Best Local Similarity 79.3%; Pred. No. 1,1e-39;

Matches 73; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 1 MYREVVVVVFMALYVOLQGSSNHEGPKVRSOSQTLERSEOOIRASSLEELRITTHSE 60

DB 1 MYGEMGMGNIIMFHYLVQGRSEHGPKVRSRSLERSBOQIRASSLEELQIAHSE 60

QY 61 DKMLMRCRLRLKSFSTMSDRSASHRSTRFAAT 92

DB 61 DKMLMRCRLRLKSFSTMSDRSASHRSTRFAAT 92

RESULT 3

US-08-915-795-8

Sequence 8, Application US/08915795

Patent No. 6235713

GENERAL INFORMATION:

APPLICANT: Marc G. ACHEN

APPLICANT: Andrew F. WILKS

APPLICANT: Steven A. STACKER

APPLICANT: Kari ALITALO

TITLE OF INVENTION: GROWTH FACTOR

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.

STREET: 1200 G Street, NW, Suite 700

CITY: Washington

STATE: DC

COUNTRY: United States of America

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/915,795

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: EVANS, Joseph D.

REGISTRATION NUMBER: 26,269

REFERENCE/DOCKET NUMBER: 1064/42983

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 628-8800

TELEFAX: (202) 628-8844

TELEX: N/A

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 358 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

TISSUE TYPE: Mouse Lung

US-08-915-795-8

Query Match 77.0%; Score 360.5; DB 4; Length 358;

Best Local Similarity 75.3%; Pred. No. 5e-38;

Matches 73; Conservative 7; Mismatches 12; Indels 5; Gaps 1;

QY 1 MYREVVVVVFMALYVOLQGSSNHEGPKV-----RSSQTLERSEOOIRASSLEELR 55

DB 1 MYGEMGMGNIIMFHYLVQGRSEHGPKVDFSRSSSMLEERSBOQIRASSLEELQ 60

QY 56 ITHSEDKMLMRCRLRLKSFSTMSDRSASHRSTRFAAT 92

DB 61 IAHSEDKMLMRCRLRLKSLASMDRSASHRSTRFAAT 97

RESULT 4

US-08-915-795-3

Sequence 3, Application US/08915795

Patent No. 6235713

GENERAL INFORMATION:

APPLICANT: Marc G. ACHEN

APPLICANT: Andrew F. WILKS

APPLICANT: Steven A. STACKER

APPLICANT: Kari ALITALO

TITLE OF INVENTION: GROWTH FACTOR

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.

STREET: 1200 G Street, NW, Suite 700

CITY: Washington

STATE: DC
COUNTRY: United States of America
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,795
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/42983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
TELEX: N/A
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 325 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
TISSUE TYPE: Human Breast
US-08-915-795-3

Query Match 66.0%; Score 309; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 1.6e-31;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 RSSQSLERSQOIRASSLEELRTHSEDMKIMRCRLKSFMSDSRSASHRSTRRA 90
DB 2 RSSQSLERSQOIRASSLEELRTHSEDMKIMRCRLKSFMSDSRSASHRSTRRA 61

QY 91 AT 92
11
DB 62 AT 63

RESULT 5
US-08-795-430-13
Sequence 13, Application US/08795430
Patent No. 6130071
GENERAL INFORMATION:
APPLICANT: Alltalo, Kari
APPLICANT: Joukov, Vladimir
TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,430
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/FI96/00427
FILING DATE: 01-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/671,573
FILING DATE: 28-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/601,132
FILING DATE: 14-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/585,895
FILING DATE: 12-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/510,133
FILING DATE: 01-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,011
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28967/33691
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 418 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-795-430-13

Query Match 20.9%; Score 98; DB 4; Length 418;
Best Local Similarity 36.8%; Pred. No. 0.00018;
Matches 21; Conservative 16; Mismatches 12; Indels 8; Gaps 2;

QY 41 EQOIRASSLEELRTHSEDMKIMRCRL-----KFTSMDSRSASHRSTRRA 91
DB 56 EQOIRASSLEELRTHSEDMKIMRCRL-----KFTSMDSRSASHRSTRRA 110

RESULT 6
US-08-795-430-11
Sequence 11, Application US/08795430
Patent No. 6130071
GENERAL INFORMATION:
APPLICANT: Alltalo, Kari
APPLICANT: Joukov, Vladimir
TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,430
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FI96/00427
FILING DATE: 01-AUG-1996
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/671,573
FILING DATE: 28-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/601,132
FILING DATE: 14-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/585,895
FILING DATE: 12-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/510,133
FILING DATE: 01-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,011
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28967/33691
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 415 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-795-430-11

Query Match 19.8%; Score 92.5; DB 4; Length 415;
Best Local Similarity 30.8%; Pred. No. 0.00091;
Matches 20; Conservative 17; Mismatches 25; Indels 3; Gaps 1;

QY 27 GPKRSSQSTLESEQIIRAASLELLRITHSEDKIMRCRLKSFMSRSRSHS 86
DB 46 GEVKAPEGKDL--EEQLRSVSSVDLSMVLXPDYKMKYQCLRKGGWOQPLTNTTGS 102
QY 87 TREPA 91
DB 103 VKFAA 107

RESULT 7
US-08-999-811-2
Sequence 2, Application US/08999811
Patent No. 5932540
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/999,811
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550
FILING DATE: 8-MAR-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: MARKOWICZ, KAREN R.
REGISTRATION NUMBER: 36,351
REFERENCE/DOCKET NUMBER: 1488.1000004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-999-811-2

Query Match 19.2%; Score 90; DB 2; Length 419;
Best Local Similarity 31.6%; Pred. No. 0.0019;
Matches 18; Conservative 19; Mismatches 12; Indels 8; Gaps 2;

QY 41 EQQIRAASLELLRITHSEDKIMRCRL-----KSFMSRSRSHSTRÉPA 91
DB 57 EEQLRSVSSVDLSMVLXPDYKMKYQCLRKGGWOHNRDQANLSR--TEETIKFAA 111

RESULT 8
US-09-042-105-2
Sequence 2, Application US/09042105
Patent No. 6040157
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,105
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550
FILING DATE: 8-MAR-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TO BE ASSIGNED
FILING DATE: 24-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: ERIC K. STEFFE
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1271 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-095-734-2

Query Match 14.5% Score 68; DB 1; Length 1271;
Best Local Similarity 43.2%; Pred. No. 5.5;
Matches 16; Conservative 5; Mismatches 16; Indels 0; Gaps 0;

QY 27 GPKRSSOSTLRSSEQIRASSLEELRLITHSDDPK 63
DB 835 GPMILSATQNGELSAQYAEASVEVELLGVAASEGMQ 871

RESULT 14
US-08-444-623-2
Sequence 2, Application US/08444623

Patent No. 5866403

GENERAL INFORMATION:

APPLICANT: Aldovini, Anna

APPLICANT: Young, Richard A.

TITLE OF INVENTION: Homologously Recombinant Slow Growing

TITLE OF INVENTION: Mycobacteria and Uses Therefor

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive

CITY: Lexington

STATE: MA

COUNTRY: USA

ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/444,623

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/095,734

FILING DATE: 22-JUL-1993

APPLICATION NUMBER: 07/711,334

FILING DATE: 06-JUN-1991

APPLICATION NUMBER: 07/367,894

FILING DATE: 19-JUN-1989

APPLICATION NUMBER: PCT/US90/03451

FILING DATE: 18-JUN-1990

APPLICATION NUMBER: PCT/US89/02962

FILING DATE: 07-JUL-1989

APPLICATION NUMBER: 07/361,944

FILING DATE: 05-JUN-1989

APPLICATION NUMBER: 07/223,089

FILING DATE: 22-JULY-1988

APPLICATION NUMBER: 07/216,390

FILING DATE: 07-JUL-1988

APPLICATION NUMBER: 07/163,546

FILING DATE: 03-MAR-1988

APPLICATION NUMBER: PCT/US88/00614

FILING DATE: 29-FEB-1988

APPLICATION NUMBER: 07/020,451

FILING DATE: 02-MAR-1987

ATTORNEY/AGENT INFORMATION:

NAME: Granahan, Patricia

REGISTRATION NUMBER: 32,227

REFERENCE/DOCKET NUMBER: WHI93-11M2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-861-6240

TELEFAX: 617-861-9540

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1271 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-444-623-2

Query Match 14.5% Score 68; DB 2; Length 1271;
Best Local Similarity 43.2%; Pred. No. 5.5;
Matches 16; Conservative 5; Mismatches 16; Indels 0; Gaps 0;

QY 27 GPKRSSOSTLRSSEQIRASSLEELRLITHSDDPK 63
DB 835 GPMILSATQNGELSAQYAEASVEVELLGVAASEGMQ 871

RESULT 15

US-08-471-869-2

Sequence 2, Application US/08471869

Patent No. 6022745

GENERAL INFORMATION:

APPLICANT: Aldovini, Anna

APPLICANT: Young, Richard A.

TITLE OF INVENTION: Homologously Recombinant Slow Growing

TITLE OF INVENTION: Mycobacteria and Uses Therefor

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive

CITY: Lexington

STATE: MA

COUNTRY: USA

ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/471,869

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/08267

FILING DATE: 22-JUL-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/095,734

FILING DATE: 22-JUL-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/711,334

FILING DATE: 06-JUN-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/367,894

FILING DATE: 19-JUN-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US90/03451

FILING DATE: 18-JUN-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US89/02962

FILING DATE: 07-JUL-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/361,944

FILING DATE: 05-JUN-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/223,089

FILING DATE: 22-JUL-1988

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/216,390

```

?      FILING DATE: 07-JUL-1988
?      PRIOR APPLICATION DATA:
?      APPLICATION NUMBER: US 07/163,546
?      FILING DATE: 03-MAR-1988
?      PRIOR APPLICATION DATA:
?      APPLICATION NUMBER: PCT/US88/00614
?      FILING DATE: 29-FEB-1988
?      PRIOR APPLICATION DATA:
?      APPLICATION NUMBER: US 07/020,451
?      FILING DATE: 02-MAR-1987
?      ATTORNEY/AGENT INFORMATION:
?      NAME: Graaham, Patricia
?      REGISTRATION NUMBER: 32,227
?      REFERENCE/DOCKET NUMBER: WHI93-11MA2
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: 617-861-6240
?      TELEFAX: 617-861-9540
?      INFORMATION FOR SEQ ID NO: 2:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 1271 amino acids
?      TYPE: amino acid
?      STRANDEDNESS: single
?      TOPOLOGY: linear
?      MOLECULE TYPE: protein
?
US-08-471-869-2

```

	Query Match	Best Local Similarity	Matches	Score	DB	Length
14.5%	43.2%	5	68	3	1271	
Conservative		5	Mismatches	16	Indels	Gaps
0	0	0	0	0	0	0

Search completed: October 17, 2001, 14:47:13
Job time: 199 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:48:42 ; Search time 78.16 Seconds
(without alignments)
89.663 Million cell updates/sec

Title: US-09-427-657-4_COPY_1_92

Perfect score: 468

Sequence: 1 MYREVVVVNFMMLYVLVQ.....SFTSMDSRSASHRTRFAAT 92

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	19.2	419	2	vascular endothelial
2	77	16.5	141	2	hypothetical prote
3	72.5	15.5	2039	2	probable calcium c
4	71	15.2	956	1	probable RNA-direc
5	69.5	14.9	225	2	ribonuclease III (
6	68	14.5	501	2	hypothetical prote
7	68	14.5	539	2	probable PPE prote
8	66.5	14.2	373	2	translation releas
9	65	13.9	561	2	hypothetical prote
10	65	13.9	2310	2	putative sensory tr
11	63.5	13.6	430	2	hypothetical prote
12	63.5	13.6	430	2	conserved hypothet
13	63	13.5	622	2	probable DNA repla
14	63	13.5	783	2	hypothetical prote
15	62.5	13.4	457	1	T-cell surface gly
16	62	13.2	288	2	RNA-binding protei
17	62	13.2	435	2	hypothetical prote
18	62	13.2	435	2	hypothetical prote
19	61.5	13.1	131	2	interleukin-13 - r
20	61.5	13.1	203	2	phospholipase C (E
21	61.5	13.1	869	2	hypothetical prote
22	61.5	13.1	1579	2	hypothetical prote
23	61	13.0	728	2	hypothetical prote
24	60.5	12.9	518	2	probable transcrip
25	60.5	12.9	737	2	hypothetical prote
26	60.5	12.9	929	2	hypothetical prote
27	60.5	12.9	959	1	rep-1 protein, for
28	60.5	12.9	1126	2	probable kinesin -
29	60	12.8	168	2	DNA mismatch repai
					hypothetical prote

ALIGNMENTS

RESULT 1

609 2 T45497
609 2 D83558
609 2 D86314
609 2 S51341
609 2 T31763
609 2 G84922
609 2 A59287
609 2 T40439
609 2 B83486
609 2 B72033
609 2 C86593
609 2 T13603
609 2 F84487
609 2 T18348
609 2 J04747
609 2 S49527

transposition prot
transport protein
hypothetical prote
SGDI protein - yea
hypothetical prote
hypothetical prote
myosin heavy chain
hypothetical prote
cobalamin biosynth
conserved hypothet
CT566 hypothetical
probable N-methyl-
probable ABC trans
probable pol polyp
adenylate cyclase
zona pellucida bin

vascular endothelial growth factor C precursor - human
N:Alternate names: FLT4 ligand DHM
C:Species: Homo sapiens (man)
C>Date: 27-Apr-1996 #sequence, revision 01-Nov-1996 #text-change 08-Oct-1999
C:Accession: S69207; S61795; S71443; S69208; G02659
R:Joukov, V.; Pajusola, K.; Kaipainen, A.; Chillov, D.; Lahtinen, I.; Kuk, E.; Saksel
EMBO J. 15, 1751, 1996
A:Title: Corrigendum: A novel vascular endothelial growth factor, VEGF-C, is a ligand
A:Reference number: S69207; MUID:96203094
A:Accession: S69207
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-419 <JOU>
A:Cross-references: EMBL:X64216; NID:g1177488; PIDN:CA63907.1; PID:e221096; PID:g118
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1995
A>Note: only a part of the translation is shown
A>Note: this is a revision to the sequence from reference S61795
R:Joukov, V.; Pajusola, K.; Kaipainen, A.; Chillov, D.; Lahtinen, I.; Kuk, E.; Saksel
EMBO J. 15, 290-298, 1996
A:Title: A novel vascular endothelial growth factor, VEGF-C, is a ligand for the FLT4
A:Reference number: S61795; MUID:96178224
A:Accession: S61795
A>Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 70-419 <JOU1>
A:Note: this sequence has been revised in reference S69207
A:Accession: S71443
A:Molecule type: protein
A:Residues: X, 104-120 <JOU2>
R:Lee, J.; Gray, A.; Yuan, J.; Luoh, S.M.; Avraham, H.; Wood, W.I.
submitted to the EMBL Data Library, December 1995
A:Description: Vascular endothelial growth factor related protein (VRP): A ligand and
A:Reference number: S69208
A:Accession: S69208
A:Molecule type: mRNA
A:Residues: 1-419 <LEE>
A:Cross-references: EMBL:U43142; NID:g1150988; PIDN:AA65214.1; PID:g1150989
R:Morris, J.C.
submitted to the EMBL Data Library, May 1996
A:Reference number: H01557
A:Accession: G02659
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-419 <MOR>
A:Cross-references: EMBL:U58111; NID:g1373426; PIDN:AA802909.1; PID:g1373427
C:Genetics:
A:Gene: GDB:VEGFC; VRP
A:Cross-references: GDB:3890883; OMIM:601528
F.1-12/Domain: signal sequence #status predicted <SIG>
F.13-102/Domain: propeptide #status predicted <PRO>

Best Local Similarity 24.4%; Pred. No. 8.2;
Matches 29; Conservative 18; Mismatches 30; Indels 42; Gaps 4;
OY 2 YRMVAVNFMVYVOLVOSSNEHGPVKRSSOSTLERSQO-----IRASS 49
| | | | | : : : : : | | | | | : : : : :
Db 634 YRLISSVSIYDQVLEKLFKAONE-----LEIMWQSPISKPGKGLSVHQADA 682
OY 50 LEBELRITHS-----EDWKLW---RCRLRLKSFSTMSDRSASHSTRF 89
| | | | | : : : : : | | | | | : : : : :
Db 683 IFPDLVKHVTVCAPAEADISGFDMWQDMLMADVEMRITVLSGFPPMMAARNRSCF 741

RESULT 5
C71721
ribonuclease III (rnc) Rpl17 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C:Accession: C71721
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark, U.
Nature 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499
A:Accession: C71721
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-225 <AND>
A:Cross-references: GB:AJ235270; GB:AJ235269; NID:g3860572; PIDN:CAAI4586.1; PID:g386068
A:Experimental source: strain Madrid E
C:Genetics:
A:Gene: rnc; Rpl17
C:Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology

Query Match 14.9%; Score 69.5; DB 2; Length 225;
Best Local Similarity 25.0%; Pred. No. 2.3;
Matches 17; Conservative 18; Mismatches 30; Indels 3; Gaps 1;
OY 15 YVOLVOSSNEHGPVKRSSOSTLERSQOIRASSLEELRITHSEDMKRLRKSF 74
| | | | | : : : : : | | | | | : : : : :
Db 98 YIIMTGEVAGC---RDNPNNTIENATEALIAIYLDSTNEITHNIEKIMAEFMKVONL 154
OY 75 TSMDSRSA 82
| | | | | : : : : :
Db 155 TDYDPKTA 162

RESULT 6
CA9930
hypothetical protein (uraA 3' region) - Mycobacterium bovis (strain BCG) (fragment)
C:Species: Mycobacterium bovis
C:Date: 19-May-1995 #sequence_revision 19-May-1995 #text_change 19-May-1995
C:Accession: CA9930
R:Aldovini, A.; Husson, R.N.; Young, R.A.
J. Bacteriol. 175, 7282-7289, 1993
A:Title: The uraA locus and homologous recombination in Mycobacterium bovis BCG.
A:Reference number: A49930; MUID:94042902
A:Accession: CA9930
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-501 <ALD>
A:Cross-references: GB:U01072

Query Match 14.5%; Score 68; DB 2; Length 501;
Best Local Similarity 43.2%; Pred. No. 8.5;
Matches 16; Conservative 5; Mismatches 16; Indels 0; Gaps 0;
OY 27 GPVKRSSOSTLERSQOIRASSLEELRITHSEDMK 63
| | | | | : : : : : | | | | | : : : : :
Db 65 GPMILSATONGELSAQYAEASVEELGVAASEGMQ 101

RESULT 7

A70899
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: A70899
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellon, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sultun, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987
A:Accession: A70899
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-539 <COL>
A:Cross-references: GB:Z80108; GB:AL123456; NID:g3256012; PIDN:CA802192.1; PID:e26555
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: ppe

Query Match 14.5%; Score 68; DB 2; Length 539;
Best Local Similarity 43.2%; Pred. No. 9.2;
Matches 16; Conservative 5; Mismatches 16; Indels 0; Gaps 0;
OY 27 GPVKRSSOSTLERSQOIRASSLEELRITHSEDMK 63
| | | | | : : : : : | | | | | : : : : :
Db 24 GPMILSATONGELSAQYAEASVEELGVAASEGMQ 60

RESULT 8
E70458
translation releasing factor RF-2 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 26-Aug-1999
C:Accession: E70458
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666
A:Accession: E70458
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-373 <AQE>
A:Cross-references: GB:AE000758; NID:g2984111; PIDN:AAC07656.1; PID:g2984119; GB:AE00
C:Genetics:
A:Gene: prfB
C:Superfamily: translation releasing factor

Query Match 14.2%; Score 66.5; DB 2; Length 373;
Best Local Similarity 28.1%; Pred. No. 8.9;
Matches 18; Conservative 12; Mismatches 19; Indels 15; Gaps 2;
OY 34 OSTLERSQOIRASSLEELRITHSEDMKLRFC-----RLRLKSFSTG-MD 78
| | | | | : : : : : | | | | | : : : : :
Db 63 EFTLRLKMLKSKVNDLEIVETSEDETWAMDEIKVEVRLRELEKTYLSEMD 122
OY 79 SRSA 82
| | | | | : : : : :
Db 123 AKNA 126

RESULT 9
T01950
hypothetical protein T2L5.4 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 24-Mar-1999
C:Accession: T01950
R:Giesel, C.; Smith, A.; Le, T.

A:Map position: 9
 A:Intons: 28/3; 59/1; 103/1; 170/3; 216/1; 259/3; 308/3; 359/2; 381/3; 478/2; 494/1; 54
 C:Superfamily: probable human DNA repair protein XRC9

Query Match 13.5%; Score 63; DB 2; Length 622;
 Best Local Similarity 29.6%; Pred. No. 39;
 Matches 21; Conservative 9; Mismatches 25; Indels 16; Gaps 2;

QY 4 EWMVNVNFMVLVQVVOGSSNEHG-PVKRSSOSTLERSDQ-----IRAA 47
 DB 70 ELTVTCNFILRLASLAGFTEDQADIDRSLEVLTEQEQGRLGRLMDSVLRAS 129

QY 48 SLEELLRIYH 58
 DB 130 CLPELLSALH 140

RESULT 14
 F84514
 hypothetical protein At2g14140 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse cross)
 C:Date: 02-Feb-2001 #sequence-revision 02-Feb-2001 #text-change 02-Feb-2001
 C:Accession: F84514
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Beilto, M.L.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 Euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-783 <STO>
 A:Cross-references: GB:AE002093; NID:g4587683; PIDN:AA05853.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g14140
 A:Map position: 2

Query Match 13.5%; Score 63; DB 2; Length 783;
 Best Local Similarity 24.7%; Pred. No. 51;
 Matches 24; Conservative 17; Mismatches 34; Indels 22; Gaps 5;

QY 17 OLVOGSSNEHGPVK--RSSOSTLERSDQIR-----AASSLELL-----RITHSE 60
 DB 664 ELTQGGSNLQTPPLPYMLSOEVEEIDLKVKKAKKILRDLLSLDEELWPDSSKWKQYE 723

QY 61 DWKLM--RCRLRLKSFSTSM--DSRSASHRSTRFAA 91
 DB 724 IMDWDEKKYKLLKATLLIHPDKLPRAHPEVKYLA 760

RESULT 15
 RMWST4
 T-cell surface glycoprotein CD4 precursor - mouse

N:Alternate names: T-cell differentiation antigen L3T4; T-cell surface antigen T4/Leu 3
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Jun-1987 #sequence-revision 30-Jun-1987 #text-change 16-Jul-1999
 C:Accession: A02110; A26038; A39893; A39955; I54564; I69018; A47642
 R:Tourville, B.; Gorman, S.D.; Field, E.H.; Hunkapiller, T.; Parnes, J.R.
 Science 234, 610-614, 1986
 A:Title: Isolation and sequence of L3T4 complementary DNA clones: expression in T cells
 A:Reference number: A02110; MUID:87018845

A:Accession: A02110
 A:Molecule type: mRNA
 A:Residues: 1-457 <TOD>
 A:Cross-references: GB:M13816; NID:g192070; PIDN:AAA37267.1; PID:g309112
 R:Littman, D.R.; Gertner, S.N.
 Nature 325, 453-455, 1987
 A:Title: Unusual intron in the immunoglobulin domain of the newly isolated murine CD4 (L
 A:Reference number: A26038; MUID:87115821
 A:Accession: A26038

A:Molecule type: mRNA
 A:Residues: 1-457 <LIT>
 A:Cross-references: GB:X04836; NID:g50353; PIDN:CAA28539.1; PID:g50354
 R:Gorman, S.D.; Tourville, B.; Parnes, J.R.
 Proc. Natl. Acad. Sci. U.S.A. 84, 7644-7648, 1987

A:Title: Structure of the mouse gene encoding CD4 and an unusual transcript in brain.
 A:Reference number: A39893; MUID:88041159
 A:Accession: A39893

A:Molecule type: DNA

A:Residues: 1-25, 'E', 27-457 <GOR>

A:Cross-references: GB:M17080; GB:J03003; NID:g192515; PIDN:AAA37402.1; PID:g387124
 R:Maddon, P.J.; Molineaux, S.M.; Maddon, D.E.; Zimmerman, K.A.; Godfrey, M.; Alt, F.W.
 Proc. Natl. Acad. Sci. U.S.A. 84, 9155-9159, 1987

A:Title: Structure and expression of the human and mouse T4 genes.
 A:Reference number: A39955; MUID:88097446

A:Accession: A39955

A:Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 25-457 <MAD>

A:Note: the cited Genbank accession number, J03564, is not in release 101.0
 R:Parnes, J.R.; Hunkapiller, T.

Immunol. Rev. 100, 109-127, 1987

A:Title: L3T4 and the immunoglobulin gene superfamily: New relationships between the
 A:Reference number: I54564; MUID:88152875

A:Accession: I54564

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-457 <RES>

A:Cross-references: GB:M36850; NID:g198670; PIDN:AAA39401.1; PID:g198671

A:Accession: I69018

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 208-318 <RE2>

A:Cross-references: GB:M36851; NID:g198672; PIDN:AAA39402.1; PID:g554183

R:Glasdon, B.J.; Tsagaratos, J.; Kitzbaum, L.; Maddox, J.; Mackay, C.R.; Brandon, M.
 Immunogenetics 33, 129-132, 1986

A:Title: The L3T4 antigen in mouse and the sheep equivalent are immunoglobulin-like.
 A:Reference number: A47642; MUID:86166594

A:Accession: A47642

A:Molecule type: protein

A:Residues: 27-43 <CLA>

C:Comment: This protein is expressed on most thymocytes, on a subset of mature T-cell
 C:Genetics:

A:Intons: 18/1; 74/1; 128/1; 207/1; 319/1; 386/1; 425/3; 448/2

C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology

C:Keywords: alternative initiators; duplication; glycoprotein; T-cell; transmembrane

F:1-26/Domain: signal sequence #status predicted <STC>

F:27-457/Product: T-cell surface glycoprotein CD4 #status experimental <MAT>

F:35-114/Domain: immunoglobulin homology <IM1>

F:139-190/Domain: immunoglobulin homology #status atypical <IM2>

F:220-301/Domain: immunoglobulin homology <IM3>

F:241-457/Product: CD4; brain-specific short form #status predicted <BRA>

F:321-372/Domain: immunoglobulin homology <IM4>

F:395-419/Domain: transmembrane #status predicted <TMN>

F:420-457/Domain: intracellular #status predicted <INT>

F:42-112, 159-188, 328-370/Disulfide bonds: #status predicted

F:187, 298, 323, 332/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.4%; Score 62.5; DB 1; Length 457;
 Best Local Similarity 29.7%; Pred. No. 31;
 Matches 22; Conservative 14; Mismatches 35; Indels 3; Gaps 3;

QY 8 VNVFMVLVQVVOGSSNE-HGPVKRSSOSTLERSDQIRAAASLEELRTTHSEDMKLMR 66
 DB 312 VNLVYKVAQNLNTLCEVGPSPKRLTLKQENQEAR-VSEDEKVVGVVAPET-GLNQ 369

QY 67 CRRLKSFSTMSDR 80
 DB 370 CLLSDEGDKYKMSDR 383

Wed Oct 17 15:32:56 2001

us-09-427-657-4_copy_1_92.rpr

Page 6

Search completed: October 17, 2001, 14:48:44
Job time: 290 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:51:47 ; Search time 42.92 Seconds

(without alignments)
73.427 Million cell updates/sec

Title: US-09-427-657-4_COPY_1_92

Perfect score: 468

Sequence: 1 MYREMYVYVNFVMLYVQLVQ.....SFTGMSDRSHSRSTRFRAT 92

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database :

SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92.5	19.8	415	1 VEGC_MOUSE	P97953 mus musculus
2	90	19.2	419	1 VEGC_HUMAN	P49767 homo sapien
3	72.5	15.5	2039	1 CCHL_YEAST	P50077 saccharomyc
4	71	15.2	956	1 RPO_SBMV	P21405 southern be
5	69.5	14.9	225	1 RNC_RICPR	O92831 rickettsia
6	66.5	14.2	373	1 RF2_AOUAE	O67695 aquilex aeo
7	63	13.5	622	1 FAGC_HUMAN	O15877 homo sapien
8	62.5	13.4	457	1 CD4_MOUSE	P06332 mus musculus
9	61.5	13.1	131	1 IL13_RAT	P42203 rattus norv
10	60.5	12.9	518	1 YL92_YEAST	P18634 saccharomyc
11	60.5	12.9	1091	1 MSH3_MOUSE	P13705 mus musculus
12	60	12.8	621	1 HEM1_AGABI	O92403 agarcus bl
13	59.5	12.7	468	1 ATPB_THIFE	P41168 thlobacillu
14	59.5	12.7	2145	1 CYAA_PODAN	O01513 podospora a
15	59	12.6	149	1 SP17_MOUSE	O62252 mus musculus
16	59	12.6	829	1 TOP1_XENLA	P41512 xenopus lae
17	59	12.6	896	1 RPOB_NENCR	P33540 neurospora
18	58.5	12.5	1783	1 Y468_MYCGE	O49460 mycoplasma
19	58.5	12.5	972	1 TOP1_DROME	P30189 drosophila
20	58	12.4	361	1 YC25_YEAST	P25377 saccharomyc
21	58	12.4	624	1 RNFB_HUMAN	O9unw22 homo sapien
22	58	12.4	655	1 SYR_HUMAN	P54136 homo sapien
23	58	12.4	905	1 CORG_SCHPO	P87140 schizosacch
24	57.5	12.3	410	1 FBW4_MOUSE	O991j2 mus musculus
25	57	12.2	110	1 RN_BACCT	P35078 bacillus ci
26	57	12.2	198	1 ATPF_MYCGA	O09543 caenorhabdi
27	57	12.2	664	1 ZAAA_CAEEL	O15431 homo sapien
28	57	12.2	976	1 SCPL_HUMAN	O997d5 homo sapien
29	57	12.2	1103	1 AKR2_HUMAN	P13539 mesocricetu
30	57	12.2	1939	1 MYSA_MESAU	P89876 1 genome po
31	56.5	12.1	320	1 MCA2_HUMAN	O13155 homo sapien
32	56.5	12.1	479	1 CBF5_CANAL	O43101 candida alb
33	56.5	12.1	479	1 CBF5_CANAL	O43101 candida alb

34	56	12.0	494	1	GPPA_ECOLI	P25552 escherichia
35	56	12.0	661	1	SYR_CRILLO	P37880 cricetus
36	56	12.0	815	1	RBW5_HUMAN	P52756 homo sapien
37	56	12.0	1403	1	YDF3_SCHPO	O10475 schizosacch
38	56	12.0	1490	1	CRK7_HUMAN	O9nyv4 homo sapien
39	56	12.0	1938	1	MYSA_MOUSE	O02566 mus musculus
40	56	12.0	1938	1	MYSA_RAT	P02563 rattus norv
41	56	12.0	1939	1	MYSA_HUMAN	P13533 homo sapien
42	55.5	11.9	274	1	TKTN_METJA	O58094 methanococc
43	55.5	11.9	457	1	CD4_RAT	P05540 rattus norv
44	55.5	11.9	474	1	CC2H_CRIFA	O01917 crithidia f
45	55.5	11.9	483	1	CBF5_YEAST	P33322 saccharomyc

ALIGNMENTS

RESULT 1
VEGC_MOUSE
ID VEGC_MOUSE STANDARD; PRT; 415 AA.
AC P97953;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR (VEGF-C) (FLT4 LIGAND)
DE (FLTR-L).
DE VEGC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE=97164697; PubMed=9012504;
RA Kulk E., Lymboussaki A., Taira S., Kaipainen A., Jeltsch M.,
RA Joukov V., Allitalo K.;
RT "VEGF-C receptor binding and pattern of expression with VEGFR-3
suggests a role in lymphatic vascular development.";
RL Development 122:3829-3837(1996).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE=97388482; PubMed=9247316;
RA Filz L.J., Morris J.C., Towler P., Long A., Burgess P., Greco R.,
RA Wang J., Gassaway R., Nickbarg E., Kovacic S., Charietta A.,
RA Giannotti J., Finerly H., Zollner R., Beler D.R., Leak L.V.,
RA Turner K.J., Wood C.R.;
RT "Characterization of murine Flt4 ligand/VEGF-C";
RL Oncogene 15:613-618(1997).
CC -!- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
CELL GROWTH.
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC
CC EMBL: U73620; AAC52984.1; -
CC EMBL: U58112; AAB46707.1; -
CC HSP: P15692; IVP.
CC MGD: MGI:109124; Vegfc.
CC InterPro: IPR000072; -
CC InterPro: IPR002400; -
CC Pfam: PF00341; PDGF; 1.
CC PRINTS: PR00438; GFCYSKNOT.
CC PROSITE: PS00249; PDGF_1; 1.
CC PROSITE: PS00278; PDGF_2; 1.

KW Mitogen: Growth factor; Glycoprotein; Signal; Repeat.
 FT SIGNAL 1 ? POTENTIAL.
 FT PROPEP 99 415 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
 FT CHAIN 271 361 4 x 24 AA TANDEN REPEATS.
 FT DOMAIN 271 294 1.
 FT REPEAT 271 294 2.
 FT REPEAT 295 318 3.
 FT REPEAT 319 342 4 (PARTIAL).
 FT REPEAT 343 361 4 (PARTIAL).
 FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 415 AA; 46471 MW; D9B3D3CECC659D6 CRC64;
 Query Match 19.8%; Score 92.5; DB 1; Length 415;
 Best Local Similarity 30.8%; Pred. No. 0.0044; Indels 3; Gaps 1;
 Matches 20; Conservative 17; Mismatches 25;
 QY 27 GPVKRSSQSLERSEQOIRASSLEELIRTHSEDMKLMRCRLKSFMSDRSASHRS 86
 DB 46 GEYKAEKQDL--EEQLRSVSSVDLMTLVPEYKMKYCKQKRGQWQNPITMTRTGDS 102
 QY 87 TRFAA 91
 DB 103 VKFAA 107
 RESULT 2
 VEGC_HUMAN STANDARD; PRT; 419 AA.
 AC P49767;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR (VEGF-C) (VASCULAR
 DE ENDOTHELIAL GROWTH FACTOR RELATED PROTEIN) (VRP) (FLT4 LIGAND) (FLT4-
 DE L).
 GN VEGFC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 OX 11
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 103-120.
 RX MEDLINE=96178224; PubMed=8617204;
 RA Joukov V., Pajusola K., Kaipainen A., Chillov D., Lahtinen I., Kuk E.,
 RA Saksela O., Kalkkinen N., Allitalo K.,
 RA "A novel vascular endothelial growth factor, VEGF-C, is a ligand for
 RA the Flt4 (VEGFR-3) and KDR (VEGFR-2) receptor tyrosine kinases.";
 RL EMBO J. 15:290-298(1996).
 RN 12
 RP ERRATUM.
 RX MEDLINE=96203094; PubMed=8612600;
 RA Joukov V., Pajusola K., Kaipainen A., Chillov D., Lahtinen I., Kuk E.,
 RA Saksela O., Kalkkinen N., Allitalo K.,
 RL EMBO J. 15:1751-1751(1996).
 RN 13
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96312526; PubMed=8700872;
 RA Lee J., Gray A., Yuan J., Luoh S.-M., Avraham H., Wood W.I.,
 RA "Vascular endothelial growth factor-related protein: a ligand and
 RA specific activator of the tyrosine kinase receptor Flt4.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:1988-1992(1996).
 RN 14
 RP SEQUENCE FROM N.A.
 RA Fitz L., Morris J.C., Towler P.S., Long A.J., Greco R.,
 RA Burgess P., Giannotti J., Charletta A., Hennessy D., Kovacic S.,
 RA Fitzgerald M., Scaltreto H., Weich N., Neben S., Finnerty H.,
 RA Zollner R., Wang J., Nickbarg E., Gassaway R., Turner K.,
 RP Wood C.R.;
 RA Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -i- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL

CC CELL GROWTH.
 CC -i- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -i- PTM: PROBABLY PROTEOLITICALLY PROCESSED IN THE C-TERMINUS.
 CC -i- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 CC -----
 CC DR EMBL; X94216; CAA63907.1; -
 CC DR EMBL; U43142; AA85214.1; -
 CC DR EMBL; U58111; AA802909.1; -
 CC DR HSP; P15692; LVPR.
 CC DR MIM; 601528; -
 CC DR InterPro; IP000072; -
 CC DR InterPro; IP002400; -
 CC DR Pfam; PF00341; PDGF_1.
 CC DR PRINTS; PR00438; GEGYSKNOT.
 CC DR PROSITE; PS00249; PDGF_1; 1.
 CC DR PROSITE; PS0278; PDGF_2; 1.
 CC KW Mitogen: Growth factor; Glycoprotein; Signal; Repeat.
 FT SIGNAL 1 ? POTENTIAL.
 FT PROPEP 102 102 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
 FT CHAIN 103 419 4 x 24 AA TANDEN REPEATS.
 FT DOMAIN 275 365 1.
 FT REPEAT 275 298 2.
 FT REPEAT 299 322 3.
 FT REPEAT 323 346 4 (PARTIAL).
 FT REPEAT 347 365 4 (PARTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 419 AA; 46883 MW; 9F598719DB3E014F CRC64;
 Query Match 19.2%; Score 90; DB 1; Length 419;
 Best Local Similarity 31.6%; Pred. No. 0.0086; Indels 8; Gaps 2;
 Matches 18; Conservative 19; Mismatches 12;
 QY 41 EQOIRASSLEELIRTHSEDMKLMRCRLK-----KSFMSDRSASHRSTRFAA 91
 DB 57 EEQLRSVSSVDLMTLVPEYKMKYCKQKRGQWQNPITMTRTGDS 111
 RESULT 3
 CCH1_YEAST STANDARD; PRT; 2039 AA.
 ID CCH1_YEAST
 AC P50077;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE PROBABLE CALCIUM-CHANNEL PROTEIN.
 GN CCH1 OR YGR217W OR G8501.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 OX 11
 RP SEQUENCE OF 1-1360 FROM N.A.
 RX STRAIN=S288C;
 RX MEDLINE=97435481; PubMed=9290212;
 RA Rieger M., Brueckner M., Schaefer M., Mueller-Aner S.,
 RA "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
 RA chromosome VII.";
 RL Yeast 13:1077-1090(1997).
 RN 12
 RP SEQUENCE OF 1184-2039 FROM N.A.
 RX STRAIN=S288C;
 RX MEDLINE=96267763; PubMed=8701610;

[illegible]

```

RT mitochondria."
RL Nature 396:133-140(1998).
CC -1- FUNCTION: DIGESTS DOUBLE-STRANDED RNA. INVOLVED IN THE PROCESSING
CC OF RIBOSOMAL RNA PRECURSORS AND OF SOME MRNAS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ENDONUCLEOLYTIC CLEAVAGE TO 5'-PHOSPHO-
CC MONESTER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: CONTAINS 1 DRBM (DOUBLE-STRANDED RNA-BINDING) DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE RNASE III FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AJ235270; CAA14586.1; -.
CC InterPro: IPR000999; -.
CC DR Pfam: PF00636; Ribonuclease_3; 1.
CC DR Pfam: PF00035; dsrm; 1.
CC DR PROSITE: PS50137; DS_RBD; 1.
CC DR PROSITE: PS00517; RNASE_3_1; 1.
CC DR PROSITE: PS50142; RNASE_3_2; 1.
CC KM Hydrolase; Nuclease; Endonuclease; RNA-binding.
CC FT DOMAIN 208 224 DRBM.
CC SO SEQUENCE 225 AA; 26006 MW; EBBE2BB2AF5646FB CRC64;

Query Match 14.9%; Score 69.5; DB 1; Length 225;
Best Local Similarity 25.0%; Pred. No. 0.85;
Matches 17; Conservative 18; Mismatches 30; Indels 3; Gaps 1;

QY 15 YVOIVGSSNENGPYKRSQSLERSEQDIRASSLEELRTTHSEDMKMLRCRLKSEF 74
DB 98 YIIMHGEFVAGC---RDNPNNIENATEALIAIYIDNSIEITHNIIEKLAEFMKVOML 154

QY 75 TSMDSRSA 82
DB 155 TDYDPKTA 162

RESULT 6
RP2_AOUAE STANDARD; PRT; 373 AA.
AC 067695;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE PEPTIDE CHAIN RELEASE FACTOR 2 (RP-2).
GN PRFB OR AO.1840.
OS Aquifex aeolicus.
CC Bacteria; Aquificales; Aquificaceae; Aquifex.
CC NCBI_TaxID=63363;
CC
CC SEQUENCE FROM N.A.
CC [1]
CC STRAIN=VF5;
CC MEDLINE=98196666; PubMed=9537320;
CC RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
CC Graham D.E., Overbeek R., Sneed M.A., Keller M., AuJay M., Huber R.,
CC Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
CC "The complete genome of the hyperthermophilic bacterium Aquifex
CC aeolicus.";
CC Nature 393:353-358(1998).
CC
CC -1- FUNCTION: PEPTIDE CHAIN RELEASE FACTOR 2 DIRECTS THE TERMINATION
CC OF TRANSLATION IN RESPONSE TO THE PEPTIDE CHAIN TERMINATION
CC CODONS UGA AND UAA (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE PROKARYOTIC AND MITOCHONDRIAL RELEASE
CC FACTORS FAMILY.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE000758; AAC07656.1; -.
CC InterPro: IPR000352; -.
CC DR Pfam: PF00472; RF-1; 1.
CC DR PROSITE: PS00745; RF_PROK_I; 1.
CC KM Protein biosynthesis.
CC SO SEQUENCE 373 AA; 43309 MW; A20146472563F326 CRC64;

Query Match 14.2%; Score 66.5; DB 1; Length 373;
Best Local Similarity 28.1%; Pred. No. 3.4;
Matches 18; Conservative 12; Mismatches 19; Indels 15; Gaps 2;

QY 34 OSTLSEEQDIRASSLEELRTTHSEDMKMLRC-----RLRLKSFYS--ND 78
DB 63 EETLKLKMLKESVVDLELVEITSEDTETWAMNDEIKVEYRTLRLELKTYSGEMD 122

QY 79 SRSA 82
DB 123 AKNA 126

RESULT 7
FACG_HUMAN STANDARD; PRT; 622 AA.
ID FACG_HUMAN
AC 015287;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE FANCONI ANEMIA GROUP G PROTEIN (FACG PROTEIN) (DNA-REPAIR PROTEIN
DE XRC9).
GN FANCG OR XRC9.
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC NCBI_TaxID=9606;
CC [1]
CC SEQUENCE FROM N.A.
CC RX MEDLINE=97404378; PubMed=9256465;
CC RA Liu N., Lamerdin J.E., Tucker J.D., Zhou Z.-Q., Walter C.A.,
CC Albaladejo J.S., Busch D.B., Thompson L.H.;
CC "The human XRC9 gene corrects chromosomal instability and mutagen
CC sensitivities in CHO UV40 cells.";
CC Proc. Natl. Acad. Sci. U.S.A. 94:9232-9237(1997).
CC [2]
CC SEQUENCE FROM N.A.
CC RX MEDLINE=99021385; PubMed=9806548;
CC RA De Winter J.P., Waisfisz Q., Koolmans M.A., Van Berkel C.G.M.,
CC Bosnoyan-Collins L., Alon N., Carreau M., Bender O., Demuth I.,
CC Schindler D., Pronk J.C., Arwert F., Hoeft H., Digweed M.,
CC Buchwald M., Joenje H.;
CC "The Fanconi anaemia group G gene FANCG is identical with XRC9.";
CC Nat. Genet. 20:281-283(1998).
CC [3]
CC SEQUENCE FROM N.A.
CC RP TISSUE=Epithelial, and Foreskin;
CC RC Lamerdin J.E., McCreedy P.M., Skowronski E., Adamson A.W.,
CC Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
CC Phan H., Velasco N., Gaines J., Dangnan L., Poundstone P.,
CC Christensen M., Georgescu A., Avila J., Liu S., Attix C., Andreise T.,
CC Trankheim M., Amico-Keller G., Coefield J., Duarte S., Lucas S.,
CC Bruce R., Thomas P., Quan G., Krommiller B., Arellano A.,
CC Montgomery M., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.O.,
CC Carraro A.V.;
CC Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC [4]

```

RP CHARACTERIZATION.
 RX MEDLINE=99303779; PubMed=10373536;
 RA Garcia-Higuera I., Kuang Y., Naf D., Masik J., D'Andrea A.D.;
 RT "Fancconi anemia proteins FANCA, FANCC, and FANCG/XRCC9 interact in a functional nuclear complex.";
 RL Mol. Cell. Biol. 19:4866-4873(1999).
 CC -1- FUNCTION: DNA REPAIR PROTEIN THAT MAY OPERATE IN A POSTREPLICATION REPAIR OR A CELL CYCLE CHECKPOINT FUNCTION. MAY BE IMPLICATED IN INTERSTRAND DNA CROSS-LINK REPAIR AND IN THE MAINTENANCE OF NORMAL CHROMOSOME STABILITY. CANDIDATE TUMOR SUPPRESSOR GENE.
 CC -1- SUBUNIT: BELONGS TO A MULTISUBUNIT COMPLEX COMPOSED OF FANCA, FANCC AND FANCG. THE COMPLEX IS NOT FOUND IN FA PATIENTS.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (MAJOR) AND CYTOPLASMIC (MINOR).
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS AND THYMUS. FOUND IN LYMPHOBLASTS.
 CC -1- DISEASE: DEFECTS IN FANCG ARE THE CAUSE OF ONE OF THE 8 COMPLEMENTATION GROUPS OF FANCONI ANEMIA (FA), AN AUTOSOMAL RECESSIVE DISORDER CHARACTERIZED BY PROGRESSIVE PANCYTOPENIA, A DIVERSE ASSORTMENT OF CONGENITAL MALFORMATIONS, AND A PREDISPOSITION TO THE DEVELOPMENT OF MALIGNANCIES. AT THE CELLULAR LEVEL, IT IS ASSOCIATED WITH HYPERSENSITIVITY TO DNA-DAMAGING AGENTS, CHROMOSOMAL INSTABILITY (INCREASED CHROMOSOME BREAKAGE), AND DEFECTIVE DNA REPAIR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U70310; AAB80802.1; -;
 DR EMBL: AJ007669; CAA07602.1; -;
 DR EMBL: AC004472; AAC07981.1; -;
 DR MIM: 602956; -;
 KW DNA repair; Nuclear protein.
 SQ SEQUENCE 622 AA; 68553 MW; 4BC7475472AC3C84 CRC64;
 Query Match 13.5%; Score 63; DB 1; Length 622;
 Best Local Similarity 29.6%; Pred. No. 15;
 Matches 21; Conservative 9; Mismatches 25; Indels 16; Gaps 2;
 OY 4 EWWYVVFEMLVQVQGSNENH-PYKRSSQSTLERSQ-----IRAA 47
 DB 70 ELTVTCNFTILASTAAGFTEDQADIQSLERLVETOEQGRLEQGLREIMDSYLRS 129
 OY 48 SSIPELLRITH 58
 DB 130 CLPELLSALH 140
 RESULT 8 *
 CD4_MOUSE STANDARD: PRT; 457 AA.
 AC P06332;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE T-CELL SURFACE GLYCOPROTEIN CD4 PRECURSOR (T-CELL SURFACE ANTIGEN T4/LEU-3) (T-CELL DIFFERENTIATION ANTIGEN L3T4).
 GN CD4
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87018645; PubMed=3094146;
 RA Tourville B., Gorman S.D., Field E.H., Hunkapiller T., Parnes J.R.;
 RT "Isolation and sequence of L3T4 complementary DNA clones: expression in T cells and brain.";

RL Science 234:610-614(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87115821; PubMed=3027575;
 RA Littman D.R., Gethner S.N.;
 RT "Unusual Intron in the immunoglobulin domain of the newly isolated murine CD4 (L3T4) gene.";
 RL Nature 325:453-455(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=88152875; PubMed=3326818;
 RA Parnes J.R., Hunkapiller T.;
 RT "L3T4 and the immunoglobulin gene superfamily: new relationships between the immune system and the nervous system.";
 RL Immunol. Rev. 100:109-127(1987).
 RN [4]
 RP SEQUENCE FROM N.A. (BRAIN FORM).
 RC TISSUE=Brain;
 RX MEDLINE=88041159; PubMed=2823269;
 RA Gorman S.D., Tourville B., Parnes J.R.;
 RT "Structure of the mouse gene encoding CD4 and an unusual transcript in brain.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:7644-7648(1987).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98112780; PubMed=9445485;
 RA Ansari-Lari M.A., Oeltgen J.C., Schwartz S., Zhang Z., Muzny D.M., Lu J., Gorrell J.H., Chnault A.C., Belmont J.W., Miller W., Gibbs R.A.;
 RT "Comparative sequence analysis of a gene-rich cluster at human chromosome 12p13 and its syntenic region in mouse chromosome 6.";
 RL Genome Res. 8:29-40(1998).
 RN [6]
 RP SEQUENCE OF 27-43.
 RX MEDLINE=86166694; PubMed=3082751;
 RA Classon B.J., Tsagaratos J., Kirsbaum L., Maddox J., McKay C.R., Brandon M., McKenzie I.F.C., Walker I.D.;
 RT "The L3T4 antigen in mouse and the sheep equivalent are immunoglobulin-like.";
 RL Immunogenetics 23:129-132(1986).
 RN [7]
 RP DISULFIDE BONDS.
 RX MEDLINE=86233454; PubMed=3086886;
 RA Classon B.J., Tsagaratos J., McKenzie I.F.C., Walker I.D.;
 RT "Partial primary structure of the T4 antigens of mouse and sheep: assignment of intrachain disulfide bonds.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:4499-4503(1986).
 CC -1- FUNCTION: ACCESSORY PROTEIN FOR MHC CLASS-II ANTIGEN/T-CELL RECEPTOR INTERACTION. MAY REGULATE T-CELL ACTIVATION.
 CC -1- SUBUNIT: ASSOCIATES WITH P56-LCK (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS ONE V-LIKE AND 3 C2-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M36850; AAA39401.1; -;
 DR EMBL: M13816; AAA37267.1; -;
 DR EMBL: X04836; CAA28539.1; -;
 DR EMBL: M36851; AAA39402.1; -;
 DR EMBL: M17080; AAA37403.1; -;
 DR EMBL: M17078; AAA37403.1; JOINED.
 DR EMBL: M17079; AAA37403.1; JOINED.
 DR EMBL: AC002397; AAC36010.1; -;
 DR PIR: A02110; RWMSTA.
 DR PIR: A26038; A26038.

DR HSPD: P01730; IMBR.
 DR MGD: MGI:88335; CD4.
 DR InterPro: IPR000973; -
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 19; 2.
 DR PRINTS: PR00692; CD4TCANTIGEN.
 DR Immunoglobulin domain; Transmembrane; glycoprotein; T-cell; MHC;
 DR Signal; Lipoprotein; Palmitate; Alternative splicing.
 FT SIGNAL 1 26
 FT CHAIN 27 457 T-CELL SURFACE GLYCOPROTEIN CD4.
 FT DOMAIN 27 394 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 395 417 POTENTIAL.
 FT DOMAIN 418 457 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 27 128 IG-LIKE V-TYPE DOMAIN.
 FT DOMAIN 129 207 IG-LIKE C2-TYPE DOMAIN.
 FT DOMAIN 208 317 IG-LIKE C2-TYPE DOMAIN.
 FT DOMAIN 318 374 IG-LIKE C2-TYPE DOMAIN.
 FT CARBOHD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHD 298 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHD 323 323 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFD 42 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFD 159 188
 FT DISULFD 328 370
 FT LIPID 418 418 PALMITATE (BY SIMILARITY).
 FT LIPID 421 421 PALMITATE (BY SIMILARITY).
 FT VAMPPLIC 1 240 MISSING (IN BRAIN-SPECIFIC ISOFORM).
 SQ SEQUENCE 457 AA; 51296 MW; 1BIDA7527CB00F33 CRC64;

Query Match 13.4%; Score 62.5; DB 1; Length 457;
 Best Local Similarity 29.7%; Pred. No. 12;
 Matches 22; Conservative 14; Mismatches 35; Indels 3; Gaps 3;

QY 8 VVPEMLVYVOLVQSSNE-HGPVKRSSOSTLERSEQIRAAASLEELRITHSEDMKLR 66
 DB 312 VLVVVKVKAQLNNTLCEVGFPTSPKRLTLKQENQEAR-VSEEDKVVAVAPET-GLMQ 369
 QY 67 CRLRKSTSMDSR 80
 DB 370 CLSSEGDVKKMDSR 383

RESULT 9
 IL13_RAT 113_RAT STANDARD: PRT; 131 AA.
 AC P42203;
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE INTERLEUKIN-13 PRECURSOR (IL-13) (T-CELL ACTIVATION PROTEIN P600).
 GN IL13 OR IL-13.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Kidney cortex;
 RX MEDLINE=94092138; PubMed=7916615;
 RA Lakkis F.G., Criet E.N.;
 RT "Cloning of rat interleukin-13 (IL-13) cDNA and analysis of IL-13
 RT gene expression in experimental glomerulonephritis.";
 RL Biochem. Biophys. Res. Commun. 197; 612-618(1993).
 CC -I- FUNCTION: CYTOKINE. INHIBITS INFLAMMATORY CYTOKINE PRODUCTION.
 CC STIMULATES WITH IL2 IN REGULATING INTERFERON-GAMMA SYNTHESIS.
 CC MAY BE CRITICAL IN REGULATING INFLAMMATORY AND IMMUNE RESPONSES
 CC (BY SIMILARITY).
 CC -I- SUBCELLULAR LOCATION: SECRETED.
 CC -I- SIMILARITY: BELONGS TO THE IL-4 / IL-13 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: L26913; AAA16478.1; -
 DR HSPD: P35225; 31TR.
 DR InterPro: IPR001325; -
 DR PROSITE: PS00838; INTERLEUKIN_4_13; 1.
 DR Cytokine; Glycoprotein; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 131 BY SIMILARITY.
 FT DISULFD 52 80 INTERLEUKIN-13.
 FT DISULFD 68 94 BY SIMILARITY.
 FT CARBOHD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 131 AA; 14093 MW; E5008CAB8DE8C201 CRC64;

Query Match 13.1%; Score 61.5; DB 1; Length 131;
 Best Local Similarity 28.4%; Pred. No. 3.6;
 Matches 27; Conservative 11; Mismatches 32; Indels 25; Gaps 4;

QY 12 MMLVYOLV-----QSSNHEGPKRSSOSTLERSEQIRAAASLEELRITHSEDMKLR-- 64
 DB 1 MALWYTAVALACLGSLAPGPEVRSSTSPVALRE-----LIEELSNITODQTSLSN 53
 QY 65 ---WRCRL-----RLKSTSMDSRSASHRSTR 88
 DB 54 SSMWMSVDLTAGGFCALSLTINISSCNAIHRTOR 88

RESULT 10
 IL13_YEAST 113_YEAST STANDARD: PRT; 518 AA.
 AC P18634;
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE HYPOTHETICAL 59.8 KDA PROTEIN IN SSRI-ATP10 INTERGENIC REGION.
 GN YLR392C OR L8084.13.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
 RA Favell A., Fulton L., Gattung S., Greco T., Kirsten J.,
 RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,
 RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
 RA Mardis E., Meneses S., Miller N., Nhan M., Pauley A., Peluso D.,
 RA Rifken L., Riles L., Taich A., Trevasaki E., Vignati D.,
 RA Wilcox L., Wohlman P., Vaudin M., Wilson R., Waterson R.;
 RL Submitted (JAN-1995) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-210 FROM N.A.
 RX MEDLINE=90277691; PubMed=2141026;
 RA Ackerman S.H., Tsagoloff A.;
 RT "ATP 10, a yeast nuclear gene required for the assembly of the
 RT mitochondrial F1-F0 complex.";
 RL J. Biol. Chem. 265; 9952-9959(1990).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>


```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U19729; AAB82352.1; -
CC DR EMBL: J05463; AAB05631.1; -
CC DR PIR: C36445; C36445.
CC SGI: S0004384; YLR392C.
CC KW Hypothetical protein.
CC SQ SEQUENCE 518 AA; 59758 MW; 65399EA2FFFC2C23 CRC64;

Query Match
Best Local Similarity 29.1%; Pred. No. 24;
Matches 23; Conservative 10; Mismatches 31; Indels 15; Gaps 3;

OY 18 LVGGSSNEHGPVRRSSOSTLERSEQOIRAASSLEELRITHSDMKLMRCRLAKS----- 73
DB 381 LSGIANGN-----YSSLKLEIDLNQKSNKPLDLNLEIRENDFGSMFCELRKDPIL 436
OY 74 ----FTSMDSRSASHRSTR 88
DB 437 RKLVFNEEDYR---HRGNR 452

RESULT 11
MSH3_MOUSE
ID MSH3_MOUSE STANDARD; PRT: 1091 AA.
AC P13705;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE DNA MISMATCH REPAIR PROTEIN MSH3 (REPAIR-3 PROTEIN) (REP-1).
GN MSH3 OR REP-3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
OY (1)
RN SEQUENCE FROM N.A.
RA Liu K., Niu L., Linton J.P., Crouse G.F.;
RL Submitted (Apr-1992) to the EMBL/Genbank/DBJ databases.
RN (2)
RN SEQUENCE FROM N.A.
RA MEDLINE=95011610; PubMed=7926796;
RA Liu K., Niu L., Linton J.P., Crouse G.F.;
RT "Characterization of the mouse Rep-3 gene: sequence similarities to
RT bacterial and yeast mismatch-repair proteins."
RL Gene 147:169-177(1994).
RN (3)
RN SEQUENCE OF 1-917 FROM N.A.
RA MEDLINE=89384567; PubMed=2674679;
RA Linton J.P., Yen J.-Y.J., Selby E., Chen Z., Chinsky J.M., Liu K.,
RA Kellems R.E., Crouse G.F.;
RT "Dual bidirectional promoters at the mouse dhfr locus: cloning and
RT characterization of two mRNA classes of the divergently transcribed
RT Rep-1 gene."
RL Mol. Cell. Biol. 9:3058-3072(1989).
CC -1- FUNCTION: NOT KNOWN. PROBABLE DNA-REPAIR PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M80360; AAA40052.1; -
CC DR EMBL: L10319; AAB60711.1; -
CC DR EMBL: L10295; AAB60711.1; JOINED.
CC DR EMBL: L10296; AAB60711.1; JOINED.
CC DR EMBL: L10297; AAB60711.1; JOINED.
CC DR EMBL: L10298; AAB60711.1; JOINED.

```

```

DR EMBL: L10299; AAB60711.1; JOINED.
DR EMBL: L10300; AAB60711.1; JOINED.
DR EMBL: L10301; AAB60711.1; JOINED.
DR EMBL: L10304; AAB60711.1; JOINED.
DR EMBL: L10305; AAB60711.1; JOINED.
DR EMBL: L10306; AAB60711.1; JOINED.
DR EMBL: L10307; AAB60711.1; JOINED.
DR EMBL: L10308; AAB60711.1; JOINED.
DR EMBL: L10309; AAB60711.1; JOINED.
DR EMBL: L10310; AAB60711.1; JOINED.
DR EMBL: L10311; AAB60711.1; JOINED.
DR EMBL: L10312; AAB60711.1; JOINED.
DR EMBL: L10313; AAB60711.1; JOINED.
DR EMBL: L10314; AAB60711.1; JOINED.
DR EMBL: L10315; AAB60711.1; JOINED.
DR EMBL: L10316; AAB60711.1; JOINED.
DR EMBL: L10318; AAB60711.1; JOINED.
DR EMBL: M24919; AAA40051.1; ALF_SEQ.
DR PIR: A30939; A30939.
DR PIR: A32495; A32495.
DR PIR: B32495; B32495.
DR MGD: MGI:109519; Msh3.
DR InterPro: IPR000432; -.
DR InterPro: IPR002863; -.
DR Pfam: PF00488; Muts_C_1.
DR Pfam: PF01624; Muts_N_1.
DR PROSITE: PS00486; DNA_MISMATCH_REPAIR_2; 1.
DR DNA repair; ATP-binding; DNA-binding.
FT NP_BIND 850 857 ATP (POTENTIAL).
FT CONFLICT 505 506 MV -> IL (IN REF. 2).
FT CONFLICT 864 864 T -> A (IN REF. 2).
SQ SEQUENCE 1091 AA; 123074 MW; 822BF422436F5D13 CRC64;

Query Match
Best Local Similarity 27.3%; Pred. No. 58;
Matches 18; Conservative 11; Mismatches 24; Indels 13; Gaps 2;

OY 17 QLVGSSNEH-----GPVRRSSOSTLERSEQOIRAASSLEELRITHSDMKLMRCRL 71
DB 46 KYEGDSRRKSLGCGPTFKARKAKVPEKEENISVASH-----HPEAKCLRPRIYL 97
OY 72 KSFTSM 77
DB 98 KSLKTL 103

RESULT 12
HEML_AGABI
ID HEML_AGABI STANDARD; PRT: 621 AA.
AC Q92403;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 5-AMINOLEVULINIC ACID SYNTHASE, MITOCHONDRIAL PRECURSOR (EC 2.3.1.37)
DE (DELTA-AMINOLEVULINATE SYNTHASE) (DELTA-ALA SYNTHETASE).
GN HEMI.
OS Agaricus bisporus (Common mushroom).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Agaricales;
OC Agaricaceae; Agaricus.
OX NCBI_TaxID=5341;
RN (1)
RN SEQUENCE FROM N.A.
RA STRAIN=D649;
RA Yague E., Mehak-zunic M., Wood D.A., Thurston C.F.;
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: SUCCINYL-COA + GLYCINE -> 5-AMINOLEVULINATE +
CC COA + CO(2).
CC -1- COTRATOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: FIRST AND RATE-LIMITING STEP IN HEME BIOSYNTHESIS.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.

```



```

FT REPEAT 983 1005 LRR 10.
FT REPEAT 1006 1027 LRR 11.
FT REPEAT 1028 1051 LRR 12.
FT REPEAT 1053 1074 LRR 13.
FT REPEAT 1075 1097 LRR 14.
FT REPEAT 1099 1120 LRR 15.
FT REPEAT 1233 1256 LRR 16.
FT REPEAT 1257 1280 LRR 17.
FT REPEAT 1282 1304 LRR 18.
FT REPEAT 1305 1328 LRR 19.
FT REPEAT 1357 1380 LRR 20.
FT REPEAT 1394 1419 LRR 21.
FT DOMAIN 1445 1710 PP2C-LIKE.
FT DOMAIN 1711 2145 CATALYTIC.
FT DOMAIN 36 41 POLY-SER.
SQ SEQUENCE 2145 AA; 237515 MW; 88E7EF6F4AC0687D CRC64;

```

```

Query Match 12.7%; Score 59.5; DB 1; Length 2145;
Best Local Similarity 26.6%; Pred. No. 1.7e+02;
Matches 17; Conservative 7; Mismatches 37; Indels 3; Gaps 1;

```

```

QY 14 LVYLVQSSNENHGPVKRSSQSTLERSQOIRASSLEELRTHSEDKMKRCLRLKS 73
DB 1988 LYPHALSGRIETHSKHEKQADLRIRPALLSPGS--ELSYEPDDIWSLWRVALRLRM 2044

```

```

QY 74 FTSM 77
DB 2045 LCSM 2048

```

```

RESULT 15

```

```

SP17_MOUSE STANDARD; PRT; 149 AA.

```

```

ID SP17_MOUSE

```

```

AC 062252:

```

```

DT 01-NOV-1997 (Rel. 35, Created)

```

```

DT 01-NOV-1997 (Rel. 35, Last sequence update)

```

```

DT 01-OCT-2000 (Rel. 40, Last annotation update)

```

```

DE SPERM SURFACE PROTEIN SP17 (SPERM AUTOANTIGENIC PROTEIN 17).

```

```

GN SPAL7 OR SP17.

```

```

OS Mus musculus (Mouse).

```

```

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

```

```

OX NCBI_TaxID=10090;

```

```

RN [1]

```

```

RP SEQUENCE FROM N.A.

```

```

RC STRAIN=BAIB/C; TISSUE=Testis;

```

```

RX MEDLINE=96039129; PubMed=7578682;

```

```

RA Kong W., Richardson R.T., Widgren E.E., O'Rand M.G.;

```

```

RT "Sequence and localization of the mouse sperm autoantigenic protein,

```

```

SP17."

```

```

RL Biol. Reprod. 53:579-590(1995).

```

```

CC -! FUNCTION: SPERM SURFACE ZONA PELLUCIDA BINDING PROTEIN. HELPS TO

```

```

CC BIND SPERMATOZOEA TO THE ZONA PELLUCIDA WITH HIGH AFFINITY. MIGHT

```

```

CC FUNCTION IN BINDING ZONA PELLUCIDA AND CARBOHYDRATES (BY

```

```

CC SIMILARITY).

```

```

CC -! SUBUNIT: HOMODIMER (BY SIMILARITY).

```

```

CC -! SUBCELLULAR LOCATION: MEMBRANE ASSOCIATED (POTENTIAL).

```

```

CC -! TISSUE SPECIFICITY: TESTIS- AND SPERM-SPECIFIC.

```

```

CC -! SIMILARITY: CONTAINS 1 IQ DOMAIN.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).

```

```

CC -----

```

```

CC EMBL; 246299; CAA86455.1;

```

```

CC MGD; MGI:1333778; SP17.

```

```

CC InterPro; IPR000048;

```

```

CC Pfam; PF00612; IQ; 1.

```

```

DR PROSITE; PS50096; IQ; 1.
KW Membrane. 112 141 IQ.
FT DOMAIN 149 AA; 17296 MW; C7E05D11D6AF0DC CRC64;
SQ SEQUENCE 149 AA; 17296 MW;

```

```

Query Match 12.6%; Score 59; DB 1; Length 149;
Best Local Similarity 35.9%; Pred. No. 8;
Matches 14; Conservative 6; Mismatches 19; Indels 0; Gaps 0;

```

```

QY 17 QLVQSSNENHGPVKRSSQSTLERSQOIRASSLEELR 55
DB 86 ELAKSSGRETEPVTPEESTEEERQEEFAAALKIOSLFR 124

```

```

Search completed: October 17, 2001, 14:51:49
Job time: 475 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:50:57 ; Search time 128.06 Seconds
(without alignments)
95.050 Million cell updates/sec

Title: US-09-427-657-4_COPY_1_92
Perfect score: 468
Sequence: 1 MYREMYVNVNFMILYVLYQ.....SFTSMDSRSASHRETRPAT 92

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP-REMBL_16:*
2: sp-archaea:*
3: sp-bacteria:*
4: sp-fungi:*
5: sp-human:*
6: sp-invertebrate:*
7: sp-mhc:*
8: sp-organellar:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-unclassified:*
13: sp-vertebrate:*
14: sp-virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	468	100.0	354	4 O43915	O43915 homo sapien
2	360.5	77.0	358	11 P97946	P97946 mus musculu
3	343.5	73.4	326	11 O35251	O35251 ratu mus norv
4	98	20.9	418	13 O57352	O57352 coturnix co
5	87	18.6	420	6 O9X550	O9X550 bos tauru
6	73.5	15.7	149	2 O68470	O68470 borrelia bu
7	73	15.6	148	2 O9S0H7	O9S0H7 borrelia bu
8	72	15.4	148	2 O06699	O06699 borrelia bu
9	72	15.4	149	2 O9S0B7	O9S0B7 borrelia bu
10	71	15.2	791	3 O9T787	O9T787 kluyveromy
11	70	15.0	608	4 O9NQX0	O9NQX0 homo sapien
12	69	14.7	66	2 O9S069	O9S069 borrelia bu
13	68	14.5	539	2 P71657	P71657 mycobacteri
14	68	14.5	539	2 O44806	O44806 borrelia bu
15	67	14.3	145	2 O44816	O44816 borrelia bu
16	67	14.0	203	2 P70443	P70443 mus musculu
17	65.5	13.9	561	10 O82601	O82601 arabidopsis
18	65	13.9	2310	3 O14002	O14002 schizosacch

20	64.5	13.8	743	11 O9JTX4	O9JTX4 mus musculu
21	64.5	13.8	769	5 O9WT12	O9WT12 drosophila
22	64.5	13.8	1088	10 O9LJ04	O9LJ04 arabidopsis
23	64.5	13.8	1584	4 O9UQ09	O9UQ09 homo sapien
24	64.5	13.8	1856	5 O9U5D7	O9U5D7 platula sta
25	64	13.7	289	13 O9DDU8	O9DDU8 gallus gall
26	64	13.7	800	10 O9LIL2	O9LIL2 arabidopsis
27	63.5	13.6	132	5 O9GXA0	O9GXA0 leishmania
28	63.5	13.6	430	2 P77302	P77302 escherichia
29	63.5	13.6	1887	3 O9HGP6	O9HGP6 schizosacch
30	63.5	13.6	1945	5 O9V768	O9V768 drosophila
31	63.5	13.6	2070	1 O9M9X3	O9M9X3 arabidopsis
32	63	13.5	288	1 O9HJ06	O9HJ06 thermoplasma
33	63	13.5	783	10 O9SIV4	O9SIV4 arabidopsis
34	63	13.5	960	5 O9NSV0	O9NSV0 caenorhabdi
35	63	13.5	978	10 O9EPY0	O9EPY0 arabidopsis
36	63	13.5	1066	5 O9NSY7	O9NSY7 caenorhabdi
37	63	13.5	2263	5 O9W0L0	O9W0L0 drosophila
38	63	13.5	2836	5 O9VSL8	O9VSL8 drosophila
39	62.5	13.4	155	11 O9WV12	O9WV12 mus musculu
40	62.5	13.4	433	11 O5S054	O5S054 mus musculu
41	62.5	13.4	457	11 O61396	O61396 mus musculu
42	62.5	13.4	733	4 O99464	O99464 homo sapien
43	62	13.2	288	4 O15815	O15815 homo sapien
44	62	13.2	288	11 O64283	O64283 mus musculu
45	62	13.2	435	2 O9Z3C0	O9Z3C0 chlamydia p

ALIGNMENTS

RESULT 1					
O43915	AD	O43915	PRELIMINARY:	PRT:	354 AA.
IC	O43915:				
DT	01-JUN-1998 (TREMBLrel. 06, Created)				
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)				
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)				
DE	GROWTH FACTOR FIGF.				
GN	FIGF OR VEGF-D.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=98140120; PubMed=9479493;				
RA	Rocchigiani M., Lestingi M., Luddi A., Orlandini M., Franco B.,				
RA	Rossi E., Ballabio A., Zuffardi O., Oliviero S.;				
RT	"Human FIGF: cloning, gene structure, and mapping to chromosome Xp22.1				
RT	between the FIGA and the GRPR genes.";				
RL	Genomics 47:207-216(1998).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=LUNG;				
RX	MEDLINE=97349118; PubMed=9205122;				
RA	Yamada Y., Nezu J., Shimane M., Hirata Y.;				
RT	"Molecular cloning of a novel vascular endothelial growth factor,				
RT	VEGF-D.";				
RL	Genomics 42:483-488(1997).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=98118549; PubMed=9435229;				
RA	Achen M.G., Jeltsch M., Kukk E., Maekinen T., Vitell A., Wilks A.F.,				
RA	Allitalo K., Stacker S.A.;				
RT	"Vascular endothelial growth factor D (VEGF-D) is a ligand for the				
RT	tyrosine kinases VEGF receptor 2 (Flk1) and VEGF receptor 3 (Flt4).";				
RL	Proc. Natl. Acad. Sci. U.S.A. 95:548-553(1998).				
DR	EMBL; Y12864; CAA73371.1; -				
DR	EMBL; Y12865; CAA73371.1; JOINED.				
DR	EMBL; Y12866; CAA73371.1; JOINED.				
DR	EMBL; Y12867; CAA73371.1; JOINED.				
DR	EMBL; Y12868; CAA73371.1; JOINED.				

Query Match	77.08;	Score 360.5;	DB 11;	Length 358;
Best Local Similarity	75.38;	Pred. No. 1.5e-33;		

RX MEDLINE=98167900; PubM

RX MEDLINE=98167900; PubMed=9435294;

RA Eichmann A., Corbel C., Jaffredo T., Breant V., Joukov V., Kumar V.,
 RA Altalo K., Le Douarin N.M.;
 RT "Avian VEGF-C: cloning, embryonic expression pattern and stimulation
 RT of the differentiation of VEGFR2-expressing endothelial cell
 RT precursors.";
 RL Development 125:743-752(1998).
 DR EMBL: Y15837; CAA57599.1; -.
 DR HSSP: P15692; IVP.
 DR InterPro: IPR002400; -.
 DR InterPro: IPR002400; -.
 DR Pfam: PF00341; PDGF_1.
 DR PRINTS: PR00438; GECYSKNQF.
 DR PRODOM: PD001629; -; 1.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS00278; PDGF_2; 1.
 DR SMART: SM00141; PDGF_1.
 KW Signal.
 FT SIGNAL.
 FT CHAIN 111 418
 SQ SEQUENCE 418 AA; 46839 MW; 099BFCC79151BF2B CRC64;

Query Match 20.9%; Score 98; DB 13; Length 418;
 Best Local Similarity 36.8%; Pred. No. 0.0029;
 Matches 21; Conservative 16; Mismatches 12; Indels 8; Gaps 2;

OY 41 EEOIRASSLEELRTTHSEDMKLMRCRL-----KSTSDNSASRSTRFAA 91
 DB 56. EEOIRASSVDELMTVLYPEYMKMFKCQLKGGWQHNRHSSDSTRSD--SLKFAA 110

RESULT 5
 O9XS50 PRELIMINARY; PRT: 420 AA.
 AC O9XS50;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HEART;
 RA Liu X., Yonekura H., Yamagishi S., Yamamoto Y., Yamamoto H.;
 RT "Structure and expression of bovine VEGF family.";
 RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB004275; BAA77687.1; -.
 DR HSSP: P15692; IVP.
 DR InterPro: IPR000072; -.
 DR Pfam: PF00341; PDGF_1.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS00278; PDGF_2; 1.
 DR SMART: SM00141; PDGF_1.
 KW Signal.
 FT SIGNAL.
 FT CHAIN 1 20
 SQ SEQUENCE 420 AA; 46681 MW; 58BA84317A3C8E2D CRC64;

Query Match 18.6%; Score 87; DB 6; Length 420;
 Best Local Similarity 43.3%; Pred. No. 0.054;
 Matches 13; Conservative 12; Mismatches 5; Indels 0; Gaps 0;

OY 41 EEOIRASSLEELRTTHSEDMKLMRCRL 70
 DB 58 EEOIRASSVDELMTVLYPEYMKMFKCQLR 87

RESULT 6

O068470 PRELIMINARY; PRT: 149 AA.
 ID O068470;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE MLP8 (2.9-8) LIPOPROTEIN (FRAGMENT).
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=297;
 RX MEDLINE=96236048; PubMed=8655511;
 RA Porcella S.F., Popova T.G., Akins D.R., Li M., Radolf J.D.,
 RA Norgard M.V.;
 RT "Borrelia burgdorferi supercoiled plasmids encode multicomplex tandem
 RT open reading frames and a lipoprotein gene family.";
 RL J. Bacteriol. 178:3293-3307(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=297;
 RA Yang X., Popova T.G., Hagman K.E., Wikel S.K., Schoeler G.B.,
 RA Calmano M.J., Radolf J.D., Norgard M.V.;
 RT "Identification, Characterization and Expression of Three New Members
 RT of the Borrelia burgdorferi Mlp (2.9) Lipoprotein Gene Family.";
 RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF046998; AAC36591.2; -.
 DR Lipoprotein; Plasmid.
 FT NON_TER 149 149
 SQ SEQUENCE 149 AA; 16958 MW; AA7B158B9872E140 CRC64;

Query Match 15.7%; Score 73.5; DB 2; Length 149;
 Best Local Similarity 28.8%; Pred. No. 0.65;
 Matches 19; Conservative 16; Mismatches 22; Indels 9; Gaps 2;

OY 7 VVAVFAMLYVOLYQG-SSNEHGVRKSSOST-----LEREEOIRASSLEELRT 57
 DB 3 IINILPCLFLMLNGCNSNDNTLKNNAOQTSRKRDLTKKEVQOERKRELLRER 62

OY 58 HSEDMK 63
 DB 63 LSDQK 68

RESULT 7
 O9S0H7 PRELIMINARY; PRT: 148 AA.
 AC O9S0H7;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE LIPOPROTEIN.
 GN BRS30
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Plasmid cp32-3.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 OX [1]
 RP SEQUENCE FROM N.A.
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Lathigra R.,
 RA White O., Dodson R., Hickey E.K., Gwinn M., Peterson J., van Vugt R.,
 RA Palmer N., Haft D., Rosa P., Stevenson B.;
 RT "A bacterial genome in flux: The twelve linear and nine circular
 RT extrachromosomal DNAs in an infectious isolate of the Lyme disease
 RT spirochete Borrelia burgdorferi.";
 RL Mol. Microbiol. 0:0-0(1999).
 DR EMBL: AE001576; AAF07449.1; -.
 KW Lipoprotein; Plasmid.
 SQ SEQUENCE 148 AA; 16731 MW; 1D406F6BC310895C CRC64;

DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE PR-POMAIN ZINC FINGER PROTEIN 6 ISOFORM A (FRAGMENT).
 GN PRDM6.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Du Y., Yang X.-H., Huang S.;
 RT "A family of novel PR-domain (PRDM) genes as candidate tumor suppressors."
 RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Deng Q.-D., Yang X.-H., Huang S.;
 RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF272898; AAF78078.1; -
 DR InterPro: IPR000822; -
 DR InterPro: IPR001214; -
 DR Pfam: PF00096; zf-C2H2; 4.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2; 2.
 DR DNA-binding; Metal-binding; Zinc-finger.
 FT NON_TER 1
 SQ SEQUENCE 608 AA; 66555 MW; 37D1D73EBAE87A2C CRC64;

Query Match 15.0%; Score 70; DB 4; Length 608;
 Best Local Similarity 27.8%; Pred. No. 7.1;
 Matches 15; Conservative 14; Mismatches 19; Indels 6; Gaps 2;

OY 24 NEHGPKRSSQSLTSESGQIRAAASLELLRLTHSE--DKMLNRRLKLSFT 75
 Db 447 DSGPISEGFNOINVKRVLASTPTSQL---HSEFSDMLMKCGCKFTFT 496

RESULT 12
 O9RR39 PRELIMINARY; PRT; 66 AA.
 AC O9RR39;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
 DE 2.9-10 LIPOPROTEIN (FRAGMENT).
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Plasmid cp32.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_Taxid=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=297;
 RX MEDLINE=96236048; PubMed=8655511;
 RA Porcella S.F., Popova T.G., Atkins D.R., Li M., Radolf J.D.,
 RA Norgard M.V.;
 RT "Borrelia burgdorferi supercoiled plasmids encode multicopy tandem
 open reading frames and a lipoprotein gene family.";
 RL J. Bacteriol. 178:3293-3307(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=297;
 RX MEDLINE=20002587; PubMed=10531261;
 RA Yang X., Popova T.G., Hagman K.E., Wikel S.K., Schoeler G.B.,
 RA Calmano M.J., Radolf J.D., Norgard M.V.;
 RT "Identification, characterization, and expression of three new members
 of the Borrelia burgdorferi MIP (2.9) lipoprotein gene family.";
 RL Infect. Immun. 67:6008-6018(1999).
 DR EMBL; AF047000; AAC36593.1; -
 DR Lipoprotein; Plasmid.
 FT NON_TER 66
 SQ SEQUENCE 66 AA; 7636 MW; 69E5D29D5B34A085 CRC64;

Query Match 14.7%; Score 69; DB 2; Length 66;
 Best Local Similarity 28.6%; Pred. No. 0.91;
 Matches 18; Conservative 15; Mismatches 22; Indels 8; Gaps 2;
 OY 7 VVNVFEMLYVOLVQG-SSNEHGPVKRSSQTLTSESGQI-----RAASLEELLRLTH 58
 Db 3 IINIFCFLILLNCSNNDNPTLKNNAOQTSRGRDLTOKEATPEKPKSRELLREKL 62

OY 59 SED 61
 Db 63 SED 65

RESULT 13
 O9S069 PRELIMINARY; PRT; 148 AA.
 AC O9S069;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
 DE 01-MAY-2000 (TREMblrel. 13, last annotation update)
 GN LIPOPROTEIN.
 GN BBL28.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Plasmid cp32-8.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_Taxid=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Lathigra R.,
 RA White O., Dodson R., Hickey E.K., Gwinn M., Peterson J., van-Vugt R.,
 RA Palmer N., Haft D., Rosa P., Stevenson B.;
 RT "A bacterial genome in flux: The twelve linear and nine circular
 extrachromosomal DNAs in an infectious isolate of the Lyme disease
 spirochete Borrelia burgdorferi.";
 RT Mol. Microbiol. 0:0-0(1999).
 DR EMBL; AE001580; AAF07641.1; -
 RN Lipoprotein; Plasmid.
 SQ SEQUENCE 148 AA; 16626 MW; F2A22BF1C60D7043 CRC64;

Query Match 14.7%; Score 69; DB 2; Length 148;
 Best Local Similarity 26.2%; Pred. No. 2.1;
 Matches 17; Conservative 18; Mismatches 22; Indels 8; Gaps 2;

OY 7 VVNVFEMLYVOLVQG-SSNEHGPVKRSSQTLTSESGQI-----RAASLEELLRLTH 58
 Db 3 IINIFCFLILLNCSNNDNPTLKNNAOQTSRGRDLTOKEATPEKPKSRELLREKL 62

OY 59 SEDMK 63
 Db 63 NDDOK 67

RESULT 14
 P71657 PRELIMINARY; PRT; 539 AA.
 AC P71657;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, last sequence update)
 DE 01-JUN-2000 (TREMblrel. 14, last annotation update)
 GN HYPOTHETICAL 55.4 KDA PROTEIN CY21B4.04.
 GN RV1387 OR MTCY21B4.04.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_Taxid=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RA Murphy L., Harris D., Barrell B.G., Rajandream M.A.;
 RL Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -!- SIMILARITY: BELONGS TO THE MTCY31.06C / MTCY251.15 / MTCY02B10.25C

CC FAMILY
 DR EMBL; 280108; CAB02192.1; -
 DR Tuberculin; Rv1387; -
 DR InterPro; IPR000030; -
 DR InterPro; IPR001899; -
 DR Pfam; PF00823; PFE.1.
 DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 8 28 POTENTIAL.
 FT TRANSMEM 63 83 POTENTIAL.
 FT TRANSMEM 117 137 POTENTIAL.
 FT TRANSMEM 279 299 POTENTIAL.
 FT TRANSMEM 312 332 POTENTIAL.
 FT TRANSMEM 347 367 POTENTIAL.
 FT TRANSMEM 368 388 POTENTIAL.
 FT TRANSMEM 391 411 POTENTIAL.
 SQ SEQUENCE 539 AA; 55403 MW; 7FE63A9BAC6FBEB5 CRC64;

Query Match 14.5%; Score 68; DB 2; Length 539;
 Best Local Similarity 43.2%; Pred. No. 11;
 Matches 16; Conservative 5; Mismatches 16; Indels 0; Gaps 0;

OY 27 GPYKRSSQSLERSEQIIRASLEELLRIHSEDMK 63
 DB 24 GPMLISATQNGELSAQYAPASAEVEELLGVASEGQ 60

RESULT 15
 044806 PRELIMINARY; PRT; 145 AA.
 AC 044806;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE LIPOPROTEIN.
 GN LP.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID-139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-297;
 RX MEDLINE-96236048; PubMed-8655511;
 RA Porcella S.F., Popova T.G., Akins D.R., Li M., Radolf J.D.,
 RA Norgard M.V.;
 RT "Borrelia burgdorferi supercoiled plasmids encode multicopy tandem
 RT open reading frames and a lipoprotein gene family.";
 RL J. Bacteriol. 178:3283-3307(1996).
 DR EMBL; U45422; AAB07830.1; -
 KW Lipoprotein.
 SQ SEQUENCE 145 AA; 16210 MW; A6F320D3B5D69185 CRC64;

Query Match 14.3%; Score 67; DB 2; Length 145;
 Best Local Similarity 24.6%; Pred. No. 3.5;
 Matches 16; Conservative 19; Mismatches 22; Indels 8; Gaps 2;

OY 7 VYVNFEMLVVQIVOG-SSNEHGPVKRSSQSLERSEQI-----RASLEELLRIH 58
 DB 3 ITNILEFCLFLMLNGSNDNDTLKNNAOOTKSRGRDILTKELTOEKPKSEELLKEKL 62
 OY 59 SEDWK 63
 DB 63 NDDOK 67

Search completed: October 17, 2001, 14:50:58
 Job time: 424 sec

XX 01-JUL-1997; 970S-0051426.
PR 23-AUG-1996; 96AU-0001825.
PR 23-AUG-1996; 96US-0023751.
PR 11-NOV-1996; 96AU-0003554.
PR 14-NOV-1996; 96US-0031097.
PR 05-FEB-1997; 97AU-0004954.
PR 10-FEB-1997; 970S-0038814.
PR 19-JUN-1997; 97AU-0007435.
XX
PA (LUDWIG-) INST. CANCER RES.
PA (UYHE-) UNIV. HELSINKI LICENSING LTD.
XX
PI Achen MG, Alltalo K, Stacker SA, Wilks AF.
XX
DR WPI; 1998-179057/16.
DR N-PSDB; AAV20806.
XX
XX New isolated vascular endothelial growth factor-D - used to develop
PT products for use in e.g. modifying angiogenesis or treating lung,
PT heart or intestinal disorders
XX
XX
PS Claim 16: Pages 57-58; 101pp; English.
XX
CC The sequence is that of human breast vascular endothelial growth factor
CC D (VEGF-D). VEGF-D can be used for e.g. acceleration of angiogenesis
CC in wound healing, tissue or organ transplantation, or to establish
CC collateral circulation in tissue infarction or arterial stenosis.
CC such as coronary artery disease, and inhibition of angiogenesis in
CC the treatment of cancer or of diabetic retinopathy. It can also be
CC used in the treatment of lung disorders to improve blood circulation
CC in the lung and/or gaseous exchange between the lungs and the blood
CC stream or to improve blood circulation to the heart and O2 gas
CC permeability in cases of cardiac insufficiency, to improve blood
CC flow and gaseous exchange in chronic obstructive airway disease,
CC or to treat malabsorptive syndromes in the intestinal tract.
CC Quantitation of VEGF-D in cancer biopsy specimens may be useful
CC as an indicator of future metastatic risk. Antagonists can be used
CC for treating e.g. conditions such as congestive heart failure,
CC involving accumulations of fluid in the lung resulting from
CC increases in vascular permeability. The products can also be used
CC for detection and diagnosis.
XX
SQ Sequence 325 AA:
XX
Query Match 100.0%; Score 897; DB 19; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.5e-70;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IIRSIQIPEEDRCSSHKKLCPIIDMLDSNKKCVLQENPLAGTEDHSHLOEPALCGPH 60
DB 173 IIRSIQIPEEDRCSSHKKLCPIIDMLDSNKKCVLQENPLAGTEDHSHLOEPALCGPH 232
QY 61 MMFEDERCVCVCTPPCKLIQHPKNCSCFECKESLETCQKHKLHPPTSCEDRCRPH 120
DB 233 mmfdeerccevccktpckdliqhpknscfeckesletccqkhklfhpdtcscedrcpht 292
QY 121 TRPCASGKTACAKHCRFPKERRAOGPHSRKNP 153
DB 293 trpcasgktacakhcrfipkexkraagpshsrknp 325
XX
RESULT 2
AAV97572
ID AAV97572 standard; Protein: 325 AA.
XX
AC AAV97572;
XX
DT 05-APR-2001 (first entry)
XX
DE Human VEGF-D protein sequence.
XX

KW Human; angiogenic protein; wound healing; vascular tissue repair;
KW peripheral arterial disease; critical limb ischaemia; coronary disease;
KW angiogenesis; tumour; inflammation; diabetic retinopathy; psoriasis;
KW rheumatoid arthritis; autoimmune disease; allergy; cancer; therapy;
KW infectious disease; neurodegeneration;
KW vascular endothelial growth factor-D; VEGF-D.
XX
OS Homo sapiens.
XX
PN MO200075163-A1.
XX
PD 14-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US14925.
XX
PR 03-JUN-1999; 99US-0137796.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Hu J, Cao L;
XX
DR WPI: 2001-071057/08.
DR N-PSDB; AAV91006.
XX
XX New nucleic acid encoding angiogenic proteins, useful e.g. for
PT promoting healing of wounds and treating peripheral arterial disease,
PT critical limb ischaemia or coronary disease -
XX
PS Claim 11: Page 226-227; 244pp; English.
XX
CC This sequence is vascular endothelial growth factor-D (VEGF-D),
CC which is an angiogenic protein of the invention. The angiogenic proteins
CC and the DNA sequences encoding them, are used to prevent, treat or
CC ameliorate disease and to detect diseases, or susceptibility, by
CC detecting mutations or the presence or amount of angiogenic protein
CC expression. Particularly they are used to stimulate wound healing,
CC growth of damaged bone and tissue, and for repair of vascular tissue,
CC especially peripheral arterial disease, critical limb ischaemia or
CC coronary disease. Antagonists of the sequences are used to inhibit
CC angiogenesis in tumours and to treat inflammation (where associated with
CC increased vascular permeability), diabetic retinopathy, rheumatoid
CC arthritis or psoriasis. Agonists are also useful for stimulating
CC (lymph)angiogenesis. The proteins are also used to identify specific
CC binding agents (potential therapeutic agents) and to raise antibodies.
CC The antibodies are useful as therapeutic (ant)agonists; for detection,
CC purification and targeting of proteins for in vivo or in vitro diagnosis
CC (including imaging) or for therapy (including when linked to e.g. a label
CC or cytotoxin); and for immunotyping of cells, e.g. for detecting minimal
CC residual disease or haematopoietic progenitor/stem cells. It is also
CC contemplated that the sequences might be useful for treating a very wide
CC range of other disorders, e.g. autoimmune diseases; allergy; cancer;
CC infectious diseases (viral, bacterial, fungal or parasitic);
CC neurodegeneration, also as chemotactic agents or for stimulating
CC regeneration of the nervous system etc.
XX
SQ Sequence 325 AA:
XX
Query Match 100.0%; Score 897; DB 22; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.5e-70;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IIRSIQIPEEDRCSSHKKLCPIIDMLDSNKKCVLQENPLAGTEDHSHLOEPALCGPH 60
DB 173 IIRSIQIPEEDRCSSHKKLCPIIDMLDSNKKCVLQENPLAGTEDHSHLOEPALCGPH 232
QY 61 MMFEDERCVCVCTPPCKLIQHPKNCSCFECKESLETCQKHKLHPPTSCEDRCRPH 120
DB 233 mmfdeerccevccktpckdliqhpknscfeckesletccqkhklfhpdtcscedrcpht 292
QY 121 TRPCASGKTACAKHCRFPKERRAOGPHSRKNP 153
DB 293 trpcasgktacakhcrfipkexkraagpshsrknp 325

```

RESULT      3
AAW49036
ID AAW49036 standard; Protein; 354 AA.
XX
AC AAW49036;
XX
DT 26-OCT-1998 (first entry)
XX
DE Human zvegf2 growth factor.
XX
KW Human zvegf2 growth factor; mitogen; fibroblast; smooth muscle cell;
KW venous stasis ulcer; diabetic ulcer; skin wound; chemotactic effect;
KW angiogenic effect; tumour; diabetic retinopathy; psoriasis; arthritis;
KW scleroderma.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..23
FT /note= "Signal peptide"
FT Peptide 24..108
FT /note= "Pro-region"
FT Binding-site 109..197
FT /note= "Receptor binding domain"
FT Region 206..256
FT /note= "Cysteine-rich domain"
FT Region 257..274
FT /note= "Baldiani ring motif"
FT Region 275..294
FT /note= "Baldiani ring motif"
FT Region 295..354
FT /note= "Cysteine-rich domain"
XX
PN WO9824811-A2.
XX
PD 11-JUN-1998.
XX
PF 20-NOV-1997; 97WO-US20888.
XX
PR 18-SEP-1997; 97US-0933455.
PR 06-DEC-1996; 96US-0759657.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Conklin DC, Gilbert T, Hart CE, Nygaard S, Sheppard PO;
DR WPI: 1998-333256/29.
DR N-PSDB; AAV32823.
XX
PT New isolated vascular endothelial growth factor - used to develop
PT products for treating e.g. wounds, burns, myocardial infarction,
PT tumours, psoriasis, arthritis, restenosis or organ transplants
XX
PS Claim 1; Pages 53-54; 77pp; English.
XX
CC The present sequence represents a human zvegf2 growth factor encoded
CC by the zvegf2 cDNA which was isolated from a human heart cDNA library.
CC zvegf2 protein in a dimeric form acts as a mitogen for fibroblasts or
CC smooth muscle cells. zvegf2 is claimed to be useful for stimulating the
CC revascularisation of tissue or the re-endothelialisation of vascular
CC tissue. zvegf2 is particularly claimed to be useful for the treatment
CC of full-thickness skin wounds, including venous stasis ulcers and
CC diabetic ulcers. The zvegf2 protein is also claimed to be useful as an
CC additive in tissue adhesives for promoting revascularisation of the
CC healing tissue. Antagonists against zvegf2 can be used to block its
CC mitogenic, chemotactic and angiogenic effects. The antagonists may
CC therefore be useful for reducing growth of solid tumours by inhibiting
CC neovascularisation of the developing tumour or by directly blocking
CC tumour cell growth, in the treatment of diabetic retinopathy, psoriasis,
CC arthritis, and scleroderma.
XX

```

```

SQ Sequence 354 AA:
Query Match 100.0%; Score 897; DB 19; Length 354;
Best Local Similarity 100.0%; Pred. No. 2,7e-70;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IIRRSIOIPEDRCNSHKRIQHPKNCSCFEKESLETCCQKRLFPDPCSCEDRCPFH 60
   |||||||
DB 202 IIRRSIQIPEDRCNSHKRIQHPKNCSCFEKESLETCCQKRLFPDPCSCEDRCPFH 261
   |||||||

QY 61 MAFEDRCCEVCCTPCPKDILQHPKNCSCFEKESLETCCQKRLFPDPCSCEDRCPFH 120
   |||||||
DB 262 mmafdrccevcctpcpkdilhphkncscfeckesletccqkrlfhpdtcscedrcpfh 321
   |||||||

QY 121 TRPCASGKTACAKHCRPEKERRAAGPHSRKNP 153
   |||||||
DB 322 trpcasgktacakhcrfpkeraagphsrknp 354
   |||||||

RESULT      4
AAW53241
ID AAW53241 standard; Protein; 354 AA.
XX
AC AAW53241;
XX
DT 03-AUG-1998 (first entry)
XX
DE Homo sapiens vascular endothelial growth factor D (VEGF-D).
XX
KW Vascular endothelial growth factor; VEGF-D; angiogenesis;
KW modification; acceleration; wound healing; tissue; organ;
KW transplants; collateral circulation; infarction; arterial stenosis;
KW coronary artery disease; inhibition; cancer; treatment;
KW diabetic retinopathy; lung disorders; blood circulation;
KW gaseous exchange; chronic obstructive airway disease;
KW intestinal malabsorptive syndrome; biopsy; metastatic risk;
KW detection; diagnosis; congestive heart failure.
XX
OS Homo sapiens.
XX
PN WO9807832-A1.
XX
PD 26-FEB-1998.
XX
PF 21-AUG-1997; 97WO-US14696.
XX
PR 01-JUL-1997; 97US-0051426.
PR 23-AUG-1996; 96AU-0001825.
PR 23-AUG-1996; 96US-0023751.
PR 11-NOV-1996; 96AU-0003554.
PR 14-NOV-1996; 96US-0031097.
PR 05-FEB-1997; 97AU-0004954.
PR 10-FEB-1997; 97US-0038814.
PR 19-JUN-1997; 97AU-0007435.
XX
PA (LUDWIG-) LUDWIG INST. CANCER RES.
PA (UYHE-) UNIV. HELSINKI LICENSING LTD.
XX
PI Achen MG, Altalo K, Stacker SA, Wilke AF;
DR WPI: 1998-179057/16.
DR N-PSDB; AAV20807.
XX
PT New isolated vascular endothelial growth factor-D - used to develop
PT products for use in e.g. modifying angiogenesis or treating lung,
PT heart or intestinal disorders
XX
PS Claim 16; Pages 60-61; 101pp; English.
XX
CC The sequence is that of human lung vascular endothelial growth factor
CC D (VEGF-D). VEGF-D can be used for e.g. acceleration of angiogenesis
CC in wound healing, tissue or organ transplantation, or to establish
XX

```

CC collateral circulation in tissue infarction or arterial stenosis,
CC such as coronary artery disease, and inhibition of angiogenesis in
CC the treatment of cancer or of diabetic retinopathy. It can also be
CC used in the treatment of lung disorders to improve blood circulation
CC in the lung and/or gaseous exchange between the lungs and the blood
CC stream or to improve blood circulation to the heart and O2 gas
CC permeability in cases of cardiac insufficiency, to improve blood
CC flow and gaseous exchange in chronic obstructive airway disease,
CC or to treat malabsorptive syndromes in the intestinal tract.
CC Quantitation of VEGF-D in cancer biopsy specimens may be useful
CC as an indicator of future metastatic risk. Antagonists can be used
CC for treating e.g. conditions such as congestive heart failure,
CC involving accumulations of fluid in the lung resulting from
CC increases in vascular permeability. The products can also be used
CC for detection and diagnosis.

SO Sequence 354 AA:

Query Match 100.0%; Score 897; DB 19; Length 354;
Best Local Similarity 100.0%; Pred. No. 2.7e-70;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IIRSIQIPEDRCSSHKKLCPIDMLWDSNKKCVLOENPLAGTEDHSHLOEPALCGPH 60
DB 202 IIRSIQIPEDRCSSHKKLCPIDMLWDSNKKCVLOENPLAGTEDHSHLOEPALCGPH 261
QY 61 MFMDEDRCVCVCTPPCKDLIOHPKNCSCFECKESLETCQKHKLPHPDTCSCEDRCRPH 120
DB 262 mfmddedrcvcvctppckdlilqhpknscfckeslelccqkhklfhpdtcscedrcrph 321
QY 121 TRPCASGKTACAKHCRPRKERRAAGPHSRKNP 153
DB 322 trpcasgktacakhcrrfprkerraaqphsrknp 354

RESULT 5
AAW44293 standard; Protein; 354 AA.
ID AAW44293 standard; Protein; 354 AA.
XX AAW44293;
AC AAW44293;
XX 22-JUN-1998 (first entry)
DT 22-JUN-1998 (first entry)
XX Human vascular endothelial growth factor D.
DE Human vascular endothelial growth factor D.
XX Human: vascular endothelial growth factor D; VEGF-D; gene therapy;
KM Inflammation; oedema.
XX Homo sapiens.
OS Homo sapiens.
XX WO9802543-A1.
PN WO9802543-A1.
XX 22-JAN-1998.
PD 22-JAN-1998.
XX 15-JUL-1997; 97WO-JP02456.
PF 15-JUL-1997; 97WO-JP02456.
XX 15-JUL-1996; 96JP-0185216.
PR 15-JUL-1996; 96JP-0185216.
XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX Hirata Y, Nezu J;
PI Hirata Y, Nezu J;
XX WPI: 1998-110591/10.
DR WPI: 1998-110591/10.
XX N-PSDB; AAV15156.
XX VEGF-D protein encoded by DNA - useful for, e.g. gene therapy and
PT treating oedema
PS Claim 1; Page 18-20; 52pp; Japanese.
XX The present sequence represents human vascular endothelial growth factor
CC D (VEGF-D). The VEGF-D protein, compounds and antibodies, which can bind

CC the protein, may be useful in, e.g. gene therapy and in treatment of
CC inflammation and oedema. Vectors, containing the VEGF-D DNA, and VEGF-D
CC DNA sequences may be used for screening for the compounds which bind to
CC the VEGF-D protein.

SO Sequence 354 AA:

Query Match 100.0%; Score 897; DB 19; Length 354;
Best Local Similarity 100.0%; Pred. No. 2.7e-70;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IIRSIQIPEDRCSSHKKLCPIDMLWDSNKKCVLOENPLAGTEDHSHLOEPALCGPH 60
DB 202 IIRSIQIPEDRCSSHKKLCPIDMLWDSNKKCVLOENPLAGTEDHSHLOEPALCGPH 261
QY 61 MFMDEDRCVCVCTPPCKDLIOHPKNCSCFECKESLETCQKHKLPHPDTCSCEDRCRPH 120
DB 262 mfmddedrcvcvctppckdlilqhpknscfckeslelccqkhklfhpdtcscedrcrph 321
QY 121 TRPCASGKTACAKHCRPRKERRAAGPHSRKNP 153
DB 322 trpcasgktacakhcrrfprkerraaqphsrknp 354

RESULT 6
AAB10649 standard; Protein; 354 AA.
ID AAB10649 standard; Protein; 354 AA.
XX AAB10649;
AC AAB10649;
XX 19-JAN-2001 (first entry)
DT 19-JAN-2001 (first entry)
XX Human VEGD protein.
DE Human VEGD protein.
XX VEGF-X; vascular endothelial growth factor; human; vulnery; cytostatic;
KM antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
KM angiogenesis regulator; vascularization regulator; cancer; psoriasis;
KM rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
KM tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
KM venous sore; diabetic ulcer; burns; skin graft growth; VEGD.
XX Homo sapiens.
OS Homo sapiens.
XX WO200037641-A2.
PN WO200037641-A2.
XX 29-JUN-2000.
PD 29-JUN-2000.
XX 21-DEC-1999; 99WO-US30503.
PF 21-DEC-1999; 99WO-US30503.
XX 22-DEC-1998; 98GB-0028377.
PR 22-DEC-1998; 98GB-0028377.
XX 18-MAR-1999; 99US-0124967.
PR 18-MAR-1999; 99US-0124967.
XX 08-NOV-1999; 99US-0164131.
PR 08-NOV-1999; 99US-0164131.
XX (JANNC) JANNC PHARM NV.
PA (JANNC) JANNC PHARM NV.
XX Gordon RD, Sprengel JJ, Yon JR, Dijkmans JH, Gosielska A;
PI Gordon RD, Sprengel JJ, Yon JR, Dijkmans JH, Gosielska A;
XX Dhanaraj SN, Xu J;
PI Dhanaraj SN, Xu J;
XX WPI: 2000-442669/38.
DR WPI: 2000-442669/38.
XX New vascular endothelial growth factor protein, useful for treating or
PT preventing diseases associated with inappropriate angiogenesis activity
PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
PS Disclosure: Fig 11; 127pp; English.
XX This invention describes a novel vascular endothelial growth factor-X
CC (VEGF-X) protein (Iia) and its encoding polynucleotide (Iia) which has
CC vulnery, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
CC antidiabetic activity and acts as an angiogenesis and vascularization
CC regulator. An antisense molecule of the invention is useful for treating
CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic

CC retinopathy by inhibiting angiogenic activity or inappropriate
CC vascularization including formation and proliferation of new blood
CC vessels, growth and development of tissues, tissue regeneration and organ
CC and tissue repair in a subject. The products of the invention are useful
CC for preparing medicaments for treating wounds such as dermal ulcers,
CC pressure sores, venous sores, diabetic ulcers and burns and to promote
CC skin graft growth, tissue repair, proliferation of new blood vessels,
CC tissue regeneration and organ repair by promoting angiogenic activity or
CC vascularization. This sequence represents the human VEGD protein used
CC to illustrate the method of the invention.

SQ Sequence 354 AA;

Query Match 100.0%; Score 897; DB 21; Length 354;
Best Local Similarity 100.0%; Pred. No. 2.7e-70;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IIRRSIQIPEDRCSSKSLCPIDMLDSNKKCVLQENPLAGTEDSHLQEPALCGPH 60
|||
DB 202 IIRRSIQIPEDRCSSKSLCPIDMLDSNKKCVLQENPLAGTEDSHLQEPALCGPH 261
|||
QY 61 MMFDEBRCVCVCTPCPKDLIQHPKNCSCFECKESLETCCQKHKLHPDTCSCEDRCRPFH 120
|||
DB 262 mmfdebrcevcvctpcpkdlqhpknscfeckesletccqkhklfhpdtcscedrcpfh 321
|||
QY 121 TRPCASGKTACAKHCRFPKERRAAGPHSRKNP 153
|||
DB 322 trpcasgktacakhcrfpkerraagphsrknp 354
|||

RESULT 7

AAB29049 standard; Protein; 354 AA.

ID AAB29049;

AC AAB29049;

XX 31-JAN-2001 (first entry)

DT 31-JAN-2001 (first entry)

XX 31-JAN-2001 (first entry)

DE Human VEGF-D protein sequence.

XX Human; Flt4; fms-like tyrosine kinase 4; lymphoedema;

XX Vascular endothelial growth factor receptor 3; VEGFR-3;

KW Milroy-None syndrome; lymphoedema praecox; VEGF-D;

XX Vascular endothelial growth factor D.

OS Homo sapiens.

XX Homo sapiens.

XX WO200058511-A1.

XX 05-OCT-2000.

XX 26-MAR-1999; 99WO-US06133.

XX 26-MAR-1999; 99WO-US06133.

XX 26-MAR-1999; 99WO-US06133.

XX (LUDW-) LUDWIG INST. CANCER RES.

XX (UYHE-) UNIV HELSINKI LICENSING LTD OY.

XX (UYPI-) UNIV PITTSBURGH.

XX Ferrell RE, Altitalo K, Finegold DN, Karkkainen M;

XX WPI: 2000-679298/66.

XX N-PSDB: AAC62407.

XX Screening a human subject for increased risk of developing a lymphatic

XX disorder, comprises assaying a nucleic acid to determine a mutation

XX altering the sequence of a vascular endothelial growth factor

XX receptor-3 -

XX Disclosure: Page 64-65; 76pp; English.

XX The present sequence is the protein sequence of the human vascular

CC endothelial growth factor D (VEGF-D). It was used to demonstrate the
CC methods of the invention, which involve the screening of individuals to
CC determine which vascular endothelial growth factor receptor 3 (VEGFR-3,
CC also known as Flt4 or fms-like tyrosine kinase 4) alleles they possess
CC and thus their likelihood of developing hereditary lymphoedema.
CC Conditions associated with lymphoedema include Milroy-None syndrome,
CC which is early onset lymphoedema and lymphoedema praecox, which is late
CC onset.

SQ Sequence 354 AA;

Query Match 100.0%; Score 897; DB 21; Length 354;
Best Local Similarity 100.0%; Pred. No. 2.7e-70;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IIRRSIQIPEDRCSSKSLCPIDMLDSNKKCVLQENPLAGTEDSHLQEPALCGPH 60
|||
DB 202 IIRRSIQIPEDRCSSKSLCPIDMLDSNKKCVLQENPLAGTEDSHLQEPALCGPH 261
|||
QY 61 MMFDEBRCVCVCTPCPKDLIQHPKNCSCFECKESLETCCQKHKLHPDTCSCEDRCRPFH 120
|||
DB 262 mmfdebrcevcvctpcpkdlqhpknscfeckesletccqkhklfhpdtcscedrcpfh 321
|||
QY 121 TRPCASGKTACAKHCRFPKERRAAGPHSRKNP 153
|||
DB 322 trpcasgktacakhcrfpkerraagphsrknp 354
|||

RESULT 8

AA70750 standard; Protein; 354 AA.

ID AA70750;

AC AA70750;

XX 17-AUG-2000 (first entry)

DT 17-AUG-2000 (first entry)

XX 17-AUG-2000 (first entry)

DE Human prepro-vascular endothelial growth factor D.

XX Human; receptor tyrosine kinase; RTK; Flt4; fms-like tyrosine kinase 4;

XX VEGFR-3; vascular endothelial growth factor receptor-3; chromosome 5q35;

KW cytosolic; tumour imaging; anti-tumour therapy; treatment; diagnosis;

KW neoplastic disease; lymphoma; carcinoma; breast; squamous cell; melanoma;

XX sarcoma; malignancy; VEGF-D; vascular endothelial growth factor D.

OS Homo sapiens.

XX Homo sapiens.

XX WO200021560-A1.

XX 20-APR-2000.

XX 08-OCT-1999; 99WO-US23525.

XX 08-OCT-1999; 99WO-US23525.

XX 09-OCT-1998; 98US-0169079.

XX (LUDW-) LUDWIG INST. CANCER RES.

XX (UYHE-) UNIV HELSINKI LICENSING LTD OY.

XX Allitalo K, Kaipainen A, Valtola R, Jussila J;

XX WPI: 2000-317850/27.

XX Treating neoplastic diseases such as lymphoma, carcinoma, melanomas

XX and sarcomas, involves administering a compound capable of inhibiting

XX binding of ligand proteins to fms-like tyrosine kinase-4 receptor -

XX Example 15-17; Page 142-143; 148pp; English.

XX The patent discloses a method to treat neoplastic disease characterised

XX by expression of fms-like tyrosine kinase 4 (Flt4) receptor (also

XX referred as vascular endothelial growth factor receptor-3, VEGFR-3) in

XX endothelial cells of blood vessels adjacent to malignant neoplasm. The

XX method involves administering a compound that inhibits binding of a

CC ligand to Flt4 thereby inhibiting Flt4 mediated proliferation of vascular
 CC endothelial cells. The compound is useful for treating neoplastic disease
 CC such as breast carcinomas, squamous cell carcinomas, lymphomas, melanomas
 CC and sarcomas. Flt4 receptor tyrosine kinase binding compounds can be used
 CC for manufacturing medicament useful for diagnostic screening, imaging and
 CC treatment of malignancies characterised by Flt4-expressing blood cells.
 CC The Flt4 gene maps to chromosomal region 5q35 and is expressed as 5.8 kb
 CC and 4.5 kb mRNAs which differ in their 3' sequences and are
 CC differentially expressed in HEL and DAMI cell lines. Flt4
 CC belongs to a subfamily of class III receptor tyrosine kinases (RTKs).
 CC It is used as a target for tumour imaging and anti-tumour therapy.
 CC The present sequence is a human prepro-vascular endothelial growth
 CC factor D (VEGF-D), a specific example of Flt4 binding compound. A
 CC recombinantly matured VEGF-D lacking residues 1-92 and 202-354
 CC retains the ability to activate VEGFR-2 and VEGFR-3 receptors and
 CC associate as non-covalently linked dimers.

SQ Sequence 354 AA:

Query Match 100.0%; Score 897; DB 21; Length 354;
 Best Local Similarity 100.0%; Pred. No. 2.7e-70;
 Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IIRRSIQIPEEDRCSHSHKLCPIIDMLDSNKKCVLQENPLAGTDSHSLQEPALCGPH 60
 |||||||
 DB 202 IIRRSIQIPEEDRCSHSHKLCPIIDMLDSNKKCVLQENPLAGTDSHSLQEPALCGPH 261
 |||||||
 QY 61 MMFEDRCEVCVCKTPCKDLIQHPKNCSCFECKESLETCCQKHKLHPDTCSEDRCPH 120
 |||||||
 DB 262 MMFEDRCEVCVCKTPCKDLIQHPKNCSCFECKESLETCCQKHKLHPDTCSEDRCPH 321
 |||||||
 DB 121 TRPCASGKTACAKHCRFPKEXKRAAGPHSRKRP 153
 |||||||
 DB 322 TRPCASGKTACAKHCRFPKEXKRAAGPHSRKRP 354
 |||||||

RESULT 9

AA70983
 ID AAY70983 standard; Protein; 354 AA.

XX AAY70983;

DT 09-AUG-2000 (first entry)

DE Human vascular endothelial growth factor (VEGF)-D protein.

XX Vascular endothelial growth factor-D; VEGF; human; re-endothelialisation;

KW vascular endothelial growth factor receptor; VEGFR; vascular trauma;

KW blood vessel; cardiovascular surgery; anti-restenosis agent; prevention;

XX restenosis; stenosis; percutaneous transluminal coronary angioplasty.

OS Homo sapiens.

XX key

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

XX WO200024412-A2.

PD 04-MAY-2000.

XX 26-OCT-1999; 99WO-US24054.

XX 26-OCT-1998; 98US-0105587.

XX (LUDM-) LUDMIG INST CANCER RES.

PA (UYHE-) UNIV HELSINKI LICENSING LTD OY.

PA (YLAH/) YLA-HERTUVALA S.

PI YLA-herituvala S, Allitalo K, Hiltunen MO, Jeltsch MM, Achen MG;

DR WPI; 2000-350584/30.

DR N-PSDB; AAD00340.

PT Preventing stenosis and restenosis in mammals using vascular

PT endothelial growth factor proteins or the nucleic acids encoding them -

PS Disclosure; Page 53-55; 61pp; English.

XX The present amino acid sequence is the complete human prepro-vascular
 CC endothelial growth factor (VEGF)-D. VEGF-D has the ability to stimulate
 CC re-endothelialisation of an injured blood vessel, without significant
 CC stimulation of smooth muscle cell proliferation. It can bind to and
 CC stimulate VEGFR-2 (vascular endothelial growth factor receptor) and/or
 CC VEGFR-3 phosphorylation in cells that express such receptors. An
 CC anti-restenosis agent comprising either a VEGF-D gene or protein is
 CC used in a method to reduce or prevent restenosis and stenosis of a blood
 CC vessel following vascular trauma e.g., cardiovascular surgery and
 CC percutaneous transluminal coronary angioplasty.

SQ Sequence 354 AA:

Query Match 100.0%; Score 897; DB 21; Length 354;
 Best Local Similarity 100.0%; Pred. No. 2.7e-70;
 Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IIRRSIQIPEEDRCSHSHKLCPIIDMLDSNKKCVLQENPLAGTDSHSLQEPALCGPH 60
 |||||||
 DB 202 IIRRSIQIPEEDRCSHSHKLCPIIDMLDSNKKCVLQENPLAGTDSHSLQEPALCGPH 261
 |||||||
 QY 61 MMFEDRCEVCVCKTPCKDLIQHPKNCSCFECKESLETCCQKHKLHPDTCSEDRCPH 120
 |||||||
 DB 262 MMFEDRCEVCVCKTPCKDLIQHPKNCSCFECKESLETCCQKHKLHPDTCSEDRCPH 321
 |||||||
 QY 121 TRPCASGKTACAKHCRFPKEXKRAAGPHSRKRP 153
 |||||||
 DB 322 TRPCASGKTACAKHCRFPKEXKRAAGPHSRKRP 354
 |||||||

RESULT 10

AA7573
 ID AAY97573 standard; Protein; 354 AA.

XX AAY97573;

DT 05-APR-2001 (first entry)

DE Human VEGF-D1 protein sequence.

XX Human; angiogenic protein; wound healing; vascular tissue repair;

KW peripheral arterial disease; critical limb ischaemia; coronary disease;

KW angiogenesis; tumour; inflammation; diabetic retinopathy; psoriasis;

KW rheumatoid arthritis; autoimmune disease; allergy; cancer; therapy;

XX infectious disease; neurodegeneration;

XX vascular endothelial growth factor-D1; VEGF-D1.

OS Homo sapiens.

XX WO200075163-A1.

XX 14-DEC-2000.

XX 01-JUN-2000; 2000MO-US14925.
 XX
 PR 03-JUN-1999; 99US-0137796.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Ruben SM, Hu J, Cao L;
 XX WPI; 2001-071057/08.
 DR N-PSDB; AAA91007.
 XX
 PT New nucleic acid encoding angiogenic proteins, useful e.g. for
 PT promoting healing of wounds and treating peripheral arterial disease.
 PT critical limb ischaemia or coronary disease -
 XX
 XX Claim 11: Page 228-229; 244pp; English.

CC This sequence is vascular endothelial growth factor-D1 (VEGF-D1),
 CC which is an angiogenic protein of the invention. The angiogenic proteins
 CC and the DNA sequences encoding them, are used to prevent, treat or
 CC ameliorate disease and to detect diseases, or susceptibility, by
 CC detecting mutations or the presence or amount of angiogenic protein
 CC expression. Particularly they are used to stimulate wound healing,
 CC growth of damaged bone and tissue, and for repair of vascular tissue,
 CC especially peripheral arterial disease, critical limb ischaemia or
 CC coronary disease. Antagonists of the sequences are used to inhibit
 CC angiogenesis in tumours and to treat inflammation (where associated with
 CC increased vascular permeability), diabetic retinopathy, rheumatoid
 CC arthritis or psoriasis. Agonists are also useful for stimulating
 CC (lymph)angiogenesis. The proteins are also used to identify specific
 CC binding agents (potential therapeutic agents) and to raise antibodies.
 CC The antibodies are useful as therapeutic (antagonists) for detection,
 CC purification and targeting of proteins for in vivo or in vitro diagnosis
 CC (including imaging) or for therapy (including when linked to e.g. a label
 CC or cytotoxin); and for immunotyping of cells, e.g. for detecting minimal
 CC residual disease or haematopoietic progenitor/stem cells. It is also
 CC contemplated that the sequences might be useful for treating a very wide
 CC range of other disorders, e.g. autoimmune diseases; allergy; cancer;
 CC infectious diseases (viral, bacterial, fungal or parasitic);
 CC neurodegeneration, also as chemotactic agents or for stimulating
 CC regeneration of the nervous system etc.

XX Sequence 354 AA:

Query Match 100.0%; Score 897; DB 22; Length 354;
 Best Local Similarity 100.0%; Pred. No. 2.7e-70;
 Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IIRSIQIPEEDRCSHSKICPTIDMLDSNKKCVLOEENPLAGTGDHSHLOEPALCGPH 60
 DB 202 IIRSIQIPEEDRCSHSKICPTIDMLDSNKKCVLOEENPLAGTGDHSHLOEPALCGPH 261
 OY 61 MAFEDRCEVCVCTPPCKDLIQHPKNCSECEKESLETCCQKHKLPHPTSCEDRCRPH 120
 DB 262 mmdedrcevcvctppckdlilqhpknscfeckesletccqkhklhptdcscedrcrph 321
 OY 121 TRPCASGKTACAKHCRFPKRAAGPHSRKNP 153
 DB 322 trpcasgktacakhcrrfpkeraagphsrkn 354

RESULT 11

ID AAB37606 standard; Protein: 354 AA.

AC AAB37606;

DT 27-FEB-2001 (first entry)

DE Human VEGF-D.

XX

KM Human; gene therapy: lymphatic disorder; hereditary lymphoedema; Flt4;
 KM vascular endothelial growth factor receptor-3; VEGFR-3; VEGF-C; VEGF-D;
 KM fms-like tyrosine kinase 4.

OS Homo sapiens.

XX CA2283470-A1.

PN 26-SEP-2000.

PD 29-SEP-1999; 99CA-2283470.

PR 26-MAR-1999; 99MO-US06133.

PR 16-AUG-1999; 99US-0375248.

XX (UYPI-) UNIV PITTSBURGH.

PA (UYHE-) UNIV HELSINKI LICENSING LTD OY.

PA (LUDW-) LUDWIG INST CANCER RES.

PI Altillo K, Ferrell RE, Finegold DN, Karkkainen M.

DR WPI; 2001-007762/02.

DR N-PSDB; AAC68954.

XX Disclosure: Pages 66-67; 99pp; English.

CC The present invention relates to a method for screening a human subject
 CC for an increased risk of developing a lymphatic disorder e.g. hereditary
 CC lymphoedema. The method comprises assaying nucleic acid of a human
 CC subject to determine a presence or an absence of a mutation altering the
 CC sequence or expression of vascular endothelial growth factor receptor-3
 CC (VEGFR-3)/fms-like tyrosine kinase 4 (Flt4) allele (see AAC68952 and
 CC AAB37604) and determining an increased risk of developing lymphatic
 CC disorder from presence or absence of the mutation. The presence of a
 CC mutation altering the encoded amino acid sequence or expression of at
 CC least 1 VEGFR-3 allele in the nucleic acid correlates with an increased
 CC risk of developing a lymphatic disorder. Treatment for hereditary
 CC lymphoedema can be provided through the administration of vascular
 CC endothelial growth factor C (VEGF-C) and vascular endothelial growth
 CC factor D VEGF-D genes (via gene therapy) and proteins. The present
 CC sequence is the protein sequence for VEGF-D.

XX Sequence 354 AA:

Query Match 100.0%; Score 897; DB 22; Length 354;
 Best Local Similarity 100.0%; Pred. No. 2.7e-70;
 Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IIRSIQIPEEDRCSHSKICPTIDMLDSNKKCVLOEENPLAGTGDHSHLOEPALCGPH 60
 DB 202 IIRSIQIPEEDRCSHSKICPTIDMLDSNKKCVLOEENPLAGTGDHSHLOEPALCGPH 261
 OY 61 MAFEDRCEVCVCTPPCKDLIQHPKNCSECEKESLETCCQKHKLPHPTSCEDRCRPH 120
 DB 262 mmdedrcevcvctppckdlilqhpknscfeckesletccqkhklhptdcscedrcrph 321
 OY 121 TRPCASGKTACAKHCRFPKRAAGPHSRKNP 153
 DB 322 trpcasgktacakhcrrfpkeraagphsrkn 354

RESULT 12

ID AAB70685 standard; Protein: 354 AA.

AC AAB70685;

DT 16-MAY-2001 (first entry)

XX

```

XX DE Human vascular endothelial growth factor D (VEGF-D) protein.
XX KW Human: vascular endothelial growth factor D; VEGF-D; angiogenic;
XX KW angiogenesis; c-fos induced growth factor; Fig1; cardiant; vasotropic;
XX KW ischaemic condition; ischaemia; vascular insufficiency;
XX KW peripheral vascular disease; coronary artery disease;
XX KW myocardial infarction.
XX OS Homo sapiens.
XX PN WO200112669-A1.
XX PD 22-FEB-2001.
XX PF 16-AUG-2000; 2000MO-IB01244.
XX PR 16-AUG-1999; 99US-0149300.
XX PA (UYSI-) UNIV SIENA.
XX PI Oliviero S;
XX DR WPI; 2001-202857/20.
XX PT Use of recombinant vascular endothelial growth factor or its
XX PT angiogenically active fragment or mutant, for inducing angiogenesis in
XX PT vivo or in vitro and for treating myocardial infarction, coronary
XX PT artery disease -
XX PS Claim 2; Page 51-52; 55pp; English.
XX CC The present invention describes a method for inducing angiogenesis in a
XX CC tissue, or area, in need of angiogenesis, in a mammal. The method
XX CC comprises administering recombinant vascular endothelial growth factor D
XX CC (VEGF-D) or its angiogenically active fragment or mutant (I). The
XX CC present sequence represents the human VEGF-D protein, which can be used
XX CC in the method of the invention. (I) has cardiant and vasotropic
XX CC activities, and is an angiogenesis inducer. The method can be used for
XX CC inducing angiogenesis in a tissue in need of angiogenesis, or in an area
XX CC in need of angiogenesis in a mammal. (I) is useful for treating various
XX CC ischaemic conditions manifested by vascular insufficiency such as
XX CC peripheral vascular disease, coronary artery disease or myocardial
XX CC infarction.
XX SO Sequence 354 AA:

Query Match          98.2%; Score 881; DB 22; Length 354;
Best Local Similarity 98.7%; Pred. No. 6,6e-69;
Matches 151; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 IIRRSIOIPEDGCHSKKICPIIDMLMDSNKKCVIOEENPLAGTEDHSLQEPALCGPH 60
   |||||||
DB 202 IIRRSIQIPEEDGCHSKKICPIIDMLMDSNKKCVIOEENPLAGTEDHSLQEPALCGPD 261
QY 61 MAFDEDRCEVCVCTPCPKDLIOHPKNCSEPECKESLETCCOKKHLPHPDSCSEDRCOPH 120
   |||||||
DB 262 mmfdecrcevcvctpcpkdlqhpknscfeckeeseetccgkxkhlhpdtscedtroph 321
QY 121 TRPCASGKTACAKHCRFPPEKRAAAGPSHRKP 153
   |||||||
DB 322 trpcasgktacakhcrfpekraaagpshrknp 354

RESULT 13
AAW14994
ID AAW14994 standard; Protein: 620 AA.
AC AAW14994;
XX 05-JUL-1997 (first entry)
XX

```

```

DE Human c-Fos induced growth factor (clone HF175 ORF2 product).
XX KW c-Fos induced growth factor; FIGF; Fos regulated gene;
XX KW proto-oncogene; lung disorder; cancer; tumour; therapy;
XX KW antibody; transgenic animal.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FH FT Misc-difference 16 /note= "residue 16 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 26 /note= "residue 26 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 29 /note= "residue 29 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 47 /note= "residue 47 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 71 /note= "residue 71 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 72 /note= "residue 72 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 76 /note= "residue 76 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 136 /note= "residue 136 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 220 /note= "residue 220 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 341 /note= "residue 341 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 344 /note= "residue 344 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 377 /note= "residue 377 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 435 /note= "residue 435 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 486 /note= "residue 486 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 497 /note= "residue 497 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 518 /note= "residue 518 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 541 /note= "residue 541 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 553 /note= "residue 553 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 557 /note= "residue 557 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 562 /note= "residue 562 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 579 /note= "residue 579 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"
FT FT Misc-difference 592 /note= "residue 592 corresponds to an in-frame
FT stop codon in reading frame 2 of HF175"

```


KW c-Fos induced growth factor; FIGF; Fos regulated gene;
 KM proto-oncogene; lung disorder; cancer; tumour; therapy;
 KM antibody; transgenic animal.

OS Mus sp.

XX
 XX Key Location/Qualifiers
 FH Region 112..164
 FT /note- "VEGF homology region"
 XX

PN W09712972-A2.

XX
 PD 10-APR-1997.

XX
 PF 30-SEP-1996; 96MO-IB01113.

XX
 PR 13-JUN-1996; 96GB-0012368.
 PR 29-SEP-1995; 95GB-0019928.

XX
 PA (UYSI-) UNIV SIENA.

XX
 PI Oliviero S;

XX
 DR WPI: 1997-226216/20.

DR N-PSDB; AAT62960.

XX
 PT Nucleotide molecule encoding c-Fos induced growth factor protein -
 PT useful in therapy, in manufacture of compositions for treatment of
 PT developmental disorders and in generation of transgenic animal
 XX
 PS Claim 3; Fig 1; 64pp; English.

XX
 CC Novel murine c-Fos induced growth factor (FIGF) (AAM14992) shows
 CC homology to the growth factor VEGF. It is encoded by the F0401
 CC gene (AAT62960) obtd. from mouse fibroblast cells. FIGF is a c-fos-
 CC dependent autocrine growth factor able to induce cell division
 CC entry and, when over-expressed, a transformed phenotype in
 CC fibroblasts. It could be implicated in tumours and development.
 CC Recombinant FIGF can be produced in transformed host (e.g. CHO)
 CC cells. It can be used to identify its receptors and in an assay
 CC for the identification of agonists and antagonists. Antibodies
 CC raised against FIGF can be used to block the function of the
 CC protein and thereby inhibit or suppress tumour growth. Transgenic
 CC animals expressing FIGF can be generated for use e.g. as models for
 CC research.
 CC
 SQ Sequence 358 AA;

Query Match 82.4%; Score 739.5; DB 18; Length 358.

Best Local Similarity 82.4%; Pred. No. 1.2e-56;

Matches 126; Conservative 9; Mismatches 17; Indels 1; Gaps 1;

OY 1 IIRRSIQIPEDRCSHSKLCPLDMLMDSNKKCVLOEENPLAGTEDHSHLOEPALCGPH 60
 ||||||| ||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 207 IIRRSIQIPEDRCSPKLCPLDMLMDSNKKCVLOEENPLAGTEDHSHLOEPALCGPH 266
 OY 61 MMFEDRCCEVCCTPPKDLIOHPKNSCEPECKESLETCCQKHKLPHPTPCSCEDRCPPH 120
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 267 mtfdeircevcvckapepgdliqhpencscfecesiesccqkhkltfhpdcscedrcpfn 326
 OY 121 TRPCASGKTACAHCRFPKRAAOGPHSRKRP 153
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 327 trtcasrkpacgkhwrfpketr-agqlysgenp 358

Search completed: October 17, 2001, 14:46:03
 Job time: 129 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:47:13 : Search time 62.93 Seconds
(without alignments)
50.061 Million cell updates/sec

Title: US-09-427-657-4_COPY_202_354
Perfect score: 897
Sequence: 1 IIRRSIQIPEDRCSHSKTL.....HCRPPKRAAGPSHRKNP 153

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTCTS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Dackfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	897	100.0	325	4 US-08-915-795-3	Sequence 3, Appl1
2	897	100.0	354	4 US-08-915-795-5	Sequence 5, Appl1
3	739.5	82.4	358	4 US-08-915-795-8	Sequence 8, Appl1
4	574	64.0	321	4 US-08-915-795-9	Sequence 9, Appl1
5	244.5	27.3	350	2 US-08-999-811-4	Sequence 4, Appl1
6	244.5	27.3	350	2 US-08-824-996-2	Sequence 2, Appl1
7	244.5	27.3	350	3 US-09-042-105-4	Sequence 4, Appl1
8	244.5	27.3	350	4 US-08-510-133A-33	Sequence 33, Appl1
9	244.5	27.3	350	4 US-08-585-895-33	Sequence 33, Appl1
10	244.5	27.3	419	2 US-08-999-811-2	Sequence 2, Appl1
11	244.5	27.3	419	3 US-09-042-105-2	Sequence 2, Appl1
12	244.5	27.3	419	3 US-09-042-105-18	Sequence 18, Appl1
13	244.5	27.3	419	4 US-08-795-430-8	Sequence 8, Appl1
14	244.5	27.3	419	4 US-08-510-133A-35	Sequence 35, Appl1
15	244.5	27.3	419	5 PCR-US96-09001-2	Sequence 2, Appl1
16	236.5	26.4	415	4 US-08-795-430-11	Sequence 11, Appl1
17	231	25.8	418	4 US-08-795-430-13	Sequence 13, Appl1
18	100	11.1	155	1 US-08-468-347-19	Sequence 19, Appl1
19	100	11.1	155	2 US-08-467-389-19	Sequence 19, Appl1
20	100	11.1	155	2 US-08-779-379-19	Sequence 19, Appl1
21	100	11.1	155	2 US-08-469-219-19	Sequence 19, Appl1
22	100	11.1	155	4 US-09-228-152-18	Sequence 18, Appl1
23	100	11.1	197	1 US-08-468-347-24	Sequence 24, Appl1
24	100	11.1	197	2 US-08-467-389-24	Sequence 24, Appl1
25	100	11.1	197	2 US-08-779-379-24	Sequence 24, Appl1
26	100	11.1	197	2 US-08-469-219-24	Sequence 24, Appl1
27	100	11.1	197	4 US-09-228-152-24	Sequence 24, Appl1

28	93.5	10.4	833	1 US-08-264-534-6	Sequence 6, Appl1
29	93.5	10.4	833	1 US-08-083-590A-2	Sequence 2, Appl1
30	93.5	10.4	833	1 US-08-465-500-6	Sequence 6, Appl1
31	93.5	10.4	833	2 US-08-346-126-6	Sequence 6, Appl1
32	93.5	10.4	833	2 US-08-346-128-6	Sequence 6, Appl1
33	93.5	10.4	833	2 US-08-532-384-2	Sequence 2, Appl1
34	93.5	10.4	833	3 US-08-893-828-6	Sequence 6, Appl1
35	93	10.4	2199	5 PCR-US95-11684-2	Sequence 2, Appl1
36	91.5	10.2	713	3 US-08-872-855-5	Sequence 5, Appl1
37	91.5	10.2	717	3 US-08-872-855-9	Sequence 9, Appl1
38	91.5	10.2	720	3 US-08-872-855-4	Sequence 4, Appl1
39	90.5	10.1	1810	5 PCR-US95-11684-4	Sequence 4, Appl1
40	89.5	10.0	799	2 US-08-525-940-23	Sequence 23, Appl1
41	89.5	10.0	799	2 US-08-976-838-23	Sequence 23, Appl1
42	89.5	10.0	881	2 US-08-525-940-21	Sequence 21, Appl1
43	89.5	10.0	881	2 US-08-976-838-21	Sequence 21, Appl1
44	89.5	10.0	915	2 US-08-525-940-18	Sequence 18, Appl1
45	89.5	10.0	915	2 US-08-976-838-18	Sequence 18, Appl1

ALIGNMENTS

RESULT 1
US-08-915-795-3
Sequence 3, Application US/08915795
Patent No. 6235713
GENERAL INFORMATION:
APPLICANT: Marc G. ACHEN
APPLICANT: Andrew F. WILKS
APPLICANT: Steven A. STRACKER
APPLICANT: Kari ALTRALO
TITLE OF INVENTION: GROWTH FACTOR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESS: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
STREET: 1200 G Street, NW, Suite 700
CITY: Washington
STATE: DC
COUNTRY: United States of America
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,795
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/42983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
TELEX: N/A
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 325 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
TISSUE TYPE: Human Breast
US-08-915-795-3
Query Match 100.0%; Score 897; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 1e-76;


```
Sequence 9, Application US/08915795
Patent No. 6235713
GENERAL INFORMATION:
APPLICANT: Marc G. ACHEN
APPLICANT: Andrew F. WILKS
APPLICANT: Steven A. STACKER
TITLE OF INVENTION: GROWTH FACTOR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeeown, Edwards & Lenahan P.L.L.C.
STREET: 1200 G Street, NW, Suite 700
CITY: Washington
STATE: DC
COUNTRY: United States of America
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,795
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/42983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
TELEX: N/A
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS: single
MOLECULE TYPE: protein
ORIGINAL SOURCE:
TISSUE TYPE: Mouse Lung
US-08-915-795-9

Query Match
Best Local Similarity 64.0%; Score 574; DB 4; Length 321;
Best Local Similarity 85.6%; Pred. No. 1.5e-46;
Matches 95; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 IIRSIQIPEDRCSHKKLCPIDMLMDSNKKCVLOENPLAGTSDHSHLOPALGPH 60
DB 202 IIRSIQIPEDRCSHKKLCPIDMLMDSNKKCVLOENPLAGTSDHSHLOPALGPH 261
QY 61 MFDEDRCEVCVCKPCPKDLIQHPKNCSCFECKSELETCCQKHKLFPHPDTC 111
DB 262 MFDEDRCEVCVCKPCPKDLIQHPKNCSCFECKSELETCCQKHKLFPHPDTC 312

RESULT 5
US-08-999-811-4
Sequence 4, Application US/08999811
Patent No. 5932540
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
```

```
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/999,811
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550
FILING DATE: 8-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: MARKOWICZ, KAREN R.
REGISTRATION NUMBER: 36,351
REFERENCE/DOCKET NUMBER: 1488.1000004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-999-811-4
```

```
Query Match
Best Local Similarity 27.3%; Score 244.5; DB 2; Length 350;
Best Local Similarity 29.8%; Pred. No. 1e-15;
Matches 57; Conservative 26; Mismatches 47; Indels 61; Gaps 8;

QY 1 IIRSI--QIPEDRCSHKKLCPIDMLMDSNKKCVLOE-----ENPLAGIED-- 47
DB 155 IIRSIPLATLPQ---CQANKRTCTPTNMMNNHICRCLAQDFMFSSDAGDSTDGFDIC 211
QY 48 --HSHLOE-----PALGPH-----MFEDRCE 69
DB 212 GPKKELDEFTCCVCACRAGLRPASCGFHKELDRNSCCQVCYCKNKLFPSCGACNRFEDTQ 271
QY 70 CVCKTPCPKDLIQHPKNCSCFECKSELETCCQKHKLFPHPDTCSEDRCPNTRPCASGKT 129
DB 272 CVCKRTCPRNQPLNPKGAC-ECTESPCKLKGKKFHHOTCSC-----YRRPCTNROK 324
QY 130 ACAKHCFRPE 140
DB 325 ACEPGFSSEE 335

RESULT 6
US-08-824-996-2
Sequence 2, Application US/08824996B
Patent No. 5935820
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: Polynucleotides Encoding Vascular Endothelial Growth
FILE REFERENCE: PF112D1
CURRENT APPLICATION NUMBER: US/08/824,996B
CURRENT FILING DATE: 1997-03-27
EARLIER APPLICATION NUMBER: 08/207,550
EARLIER FILING DATE: 1994-03-08
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 350
```

TYPE: PRT
ORGANISM: Homo sapiens
US-08-824-996-2

Query Match 27.3%; Score 244.5; DB 2; Length 350;

Best Local Similarity 29.8%; Pred. No. 1e-15;
Matches 57; Conservative 26; Mismatches 47; Indels 61; Gaps 8;

1 IIRSI--QIEEDRCSHKKLCPIDMLDMSNCKVLOE-----ENPLAGTED-- 47
155 IIRSLPLATLPQ---CQANKTCPTNYMNNHICRLAODEFMFSSDAGDSTGCFHDIC 211
48 --HSHLOE-----PALCGPH-----MMFEDRCE 69
212 GPNKLEDEETCCQVCRAGLRPASCGPHKELDRNSCQVCCKNKLFPSCGANREDEMTQ 271
70 CVCRTPCPKDLIOHPKNSCFECKESLETCCQKHLFHPDTCSCEDRCFHTRPCASGKT 129
272 CVCRTCPRNQPLNPGKAC-ECTESPQKCLKGKKFHQTSC-----YRRPCTNRQK 324
130 ACAKHCRFPKE 140
325 ACEPGFSYSEE 335

RESULT 7

US-09-042-105-4
Sequence 4, Application US/09042105
Patent No. 6040157
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,105
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550
FILING DATE: 8-MAR-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TO BE ASSIGNED
FILING DATE: 24-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: ERIC K. STEFFE
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ. ID NO: 4:
SEQUENCE CHARACTERISTICS:

LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-042-105-4

Query Match 27.3%; Score 244.5; DB 3; Length 350;

Best Local Similarity 29.8%; Pred. No. 1e-15;
Matches 57; Conservative 26; Mismatches 47; Indels 61; Gaps 8;

1 IIRSI--QIEEDRCSHKKLCPIDMLDMSNCKVLOE-----ENPLAGTED-- 47
155 IIRSLPLATLPQ---CQANKTCPTNYMNNHICRLAODEFMFSSDAGDSTGCFHDIC 211
48 --HSHLOE-----PALCGPH-----MMFEDRCE 69
212 GPNKLEDEETCCQVCRAGLRPASCGPHKELDRNSCQVCCKNKLFPSCGANREDEMTQ 271
70 CVCRTPCPKDLIOHPKNSCFECKESLETCCQKHLFHPDTCSCEDRCFHTRPCASGKT 129
272 CVCRTCPRNQPLNPGKAC-ECTESPQKCLKGKKFHQTSC-----YRRPCTNRQK 324
130 ACAKHCRFPKE 140
325 ACEPGFSYSEE 335

RESULT 8

US-08-510-133A-33
Sequence 33, Application US/08510133A
Patent No. 6221839
GENERAL INFORMATION:
APPLICANT: Alltalo, Karl
APPLICANT: Joukov, Vladimir
TITLE OF INVENTION: Receptor Ligand
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/510,133A
FILING DATE: 01-Aug-1995
CLASSIFICATION: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28113/32863
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-08-510-133A-33

Query Match 27.3%; Score 244.5; DB 4; Length 350;
Best Local Similarity 29.8%; Pred. No. 1e-15;


```

QY      70  CVCCTPCPKDIDJHPKNCSCSECEKESJLTCQCKIKLHPDTCSEDECPHTRPCASGKT 129
      ||| ||| : : | | | | | : | | | | | : | | | :
Db      272  CVCRTCPDNPMLPGKCAC-ECTESPCKLKGKKEFHQTCSC-----YRRPCTNRQK 324
QY      130  ACAKHCRFPKE 140
      || : : |
Db      325  ACERGFYSSE 335

RESULT 10
US-08-999-811-2
; Sequence 2, Application US/08999811
; Patent No. 5932540
;
; GENERAL INFORMATION:
;
; APPLICANT: HU, JING-SHAN
;
; APPLICANT: ROSEN, CRAIG A.
;
; APPLICANT: CAO, LIANG

```

```

RESULT 10
US-08-999-811-2
; Sequence 2, Application US/08999811
; Patent No. 5932540
;
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: CAO, LIANG
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 15
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/999,811
; FILING DATE: HEREWITH
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/207,550
; FILING DATE: 8-MAR-1994
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,968
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: MARKOWICZ, KAREN R.
; REGISTRATION NUMBER: 36,351
; REFERENCE/DOCKET NUMBER: 1488-1000004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-999-811-2

```

	Query Match	27.3%	Score 244.5	DB 2	Length 419;
	Best Local Similarity	29.8%	Pred. No. 1.2e-15;		
	Matches 57; Conservative	26;	Mismatches 47;	Indels 61;	Gaps 8
Oy	1 IIRRSI--OIPEEDRSHSKLCPIDMLDMSNKKCVLYE-----ENPLAGTED--	47			
	: :				
Dd	224 IIRSLPALLPP--CGANKTCPTNYMMNNHICRLAIEDPMFSSADGSDTIDGFHDIC	280			
Oy	48 --HSHLQE-----PALCGPH-----MMFDDEGE	69			
	: :				
Dd	281 GPKKELDETCCQVCRCAGLRPASCSGHKELDNRNSQCVCXKNKLFFPSQGANGREDEWTCQ	340			
Oy	70 CVCATCPCPDIDLQHAKNSCFECKSLSLTCCQAKKHLLHPDTSCSEDCBCHPTTPCASGKT	129			
	: :				
Dd	341 CVCKRKPCPNODLPNGKAC--ECESPKCLCLKKKKHNDHCSC-----YRPCTTNROK	393			

OY 130 ACAHCRFPKE 140
DB 394 ACEGFSYSSE 404

RESULT 11

US-09-042-105-2
; Sequence 2, Application US/09042105
; Patent No. 6040157
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,105
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/207,550
; FILING DATE: 8-MAR-1994
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,968
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: TO BE ASSIGNED
; FILING DATE: 24-DEC-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: ERIC K. STEFFE
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-042-105-2

Query Match 27.3%; Score 244.5; DB 3; Length 419;
Best Local Similarity 29.8%; Pred. No. 1.2e-15;
Matches 57; Conservative 26; Mismatches 47; Indels 61; Gaps 8;

OY 1 IIRRSI--OIPEDRCSHKKLCIPIDMLDSNCKCKVLOE-----ENPLAGTED-- 47
DB 224 IIRSLPATLPQ---CQANKTCPTNYMMNNHICRLAOEDFMFSSDAGDSTDFHIC 280
OY 48 --HSHLOE-----PALCGPH-----MMPEDRCE 69
DB 281 GPNKLEDETCQCVCRAGLRPASCGPHKELDRNSCQVCVCKNKLFPSCGANREFDENICQ 340
OY 70 CVKTCGPDILQHPKNCSCFECKESLETCQCKHKLPHPDTCSCGDRCPHRRPCASGKT 129

DB 341 CVCKTCGPNQINPGKAC-ECTESPQKLLKGRKHHQTCSC-----YRRPCTNRQK 393
OY 130 ACAHCRFPKE 140
DB 394 ACEGFSYSSE 404

RESULT 12

US-09-042-105-18
; Sequence 18, Application US/09042105
; Patent No. 6040157
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: CAO, LIANG
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,105
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/207,550
; FILING DATE: 8-MAR-1994
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,968
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: TO BE ASSIGNED
; FILING DATE: 24-DEC-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: ERIC K. STEFFE
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-042-105-18

Query Match 27.3%; Score 244.5; DB 3; Length 419;
Best Local Similarity 29.8%; Pred. No. 1.2e-15;
Matches 57; Conservative 26; Mismatches 47; Indels 61; Gaps 8;

OY 1 IIRRSI--OIPEDRCSHKKLCIPIDMLDSNCKCKVLOE-----ENPLAGTED-- 47
DB 224 IIRSLPATLPQ---CQANKTCPTNYMMNNHICRLAOEDFMFSSDAGDSTDFHIC 280
OY 48 --HSHLOE-----PALCGPH-----MMPEDRCE 69
DB 281 GPNKLEDETCQCVCRAGLRPASCGPHKELDRNSCQVCVCKNKLFPSCGANREFDENICQ 340

```
OY      70   CVCSTPCPKDIDIONPKNSCEFEKESLETCCQKHLYHPDTSCSEDBRCRPHTRCAGAKT    129  
         |||  ||| : | : | | | | | : | : | | | | | | | | | | | | : | | :  
Db       341  CVCKRTPCRBNQLPMDKCAC-ECIESPQKKLLKCKKNHNTGSC-----YRRCTIRNOK    393
```

```
OY      130  ACAKHNCPPE     140  
         ||  : : |  
Db       394  ACEPFGEYSIEE    404
```

```

RESULT 13
US-08-795-430-8
: Sequence 8, Application US/08795430
: Patent No. 6130071
:
GENERAL INFORMATION:
: APPLICANT: Allitalo, Karl
: APPLICANT: Joukov, Vladimir
: TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
: TITLE OF INVENTION: Protein and Gene, Mutants Thereof, and Uses Thereof
: NUMBER OF SEQUENCES: 57
:
CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 6300 Sears Tower, 233 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606-6402
:
COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
:
CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/795,430
: FILING DATE:
: CLASSIFICATION: 435
:
PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/FI96/00427
: FILING DATE: 01-AUG-1996
:
PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/671,573
: FILING DATE: 28-JUN-1996
:
PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/601,132
: FILING DATE: 14-FEB-1996
:
PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/585,895
: FILING DATE: 12-JAN-1996
:
PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/510,133
: FILING DATE: 01-AUG-1995
:
PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/340,011
: FILING DATE: 14-NOV-1994
:
ATTORNEY/AGENT INFORMATION:
: NAME: Gass, David A.
: REGISTRATION NUMBER: 38,153
: REFERENCE/DOCKET NUMBER: 28967/33691
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312/474-6300
: TELEFAX: 312/474-0448
: TELEX: 25-3856
:
INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 419 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
:
MOLECULE TYPE: protein
US-08-795-430-8

```

Query Match	27.38;	Score 244.5;	DB 4;	Length 419;
Best local Similarity	29.88;	Pred. No. 1.2e-15;		

	Matches	57;	Conservative	26;	Mismatches	47;	Indels	61;	Gaps	8;
Oy	1	IIRRSI--QIPEDDCSHKICLPIDMLDMSKKCVLOE-----ENPLAGED--	47							
		: : : : : : : : : : : : :								
Dd	224	IIRRSI PATL PQ---CQAANKTCPTNYMNMNHICRCLADEDFMFSSDAGDDSTDEGHFHC	280							
Oy	48	--HSHLOE-----PALCGPH-----WMFDRCCE	69							
		: : : : : : : : : : : : :								
Dd	281	GPNNKEIDETTCQCVCAGLRAPASCGPFHKRLDINSOCVCYCKNKLFPSCGACANREFDINTCO	340							
Oy	70	CYCCKTPPCRDLLIQHPKNCSCFECKSLSTCCOKIKILFIHPDTSCIEDRCPFHTRPCASGKT	129							
		: : : : : : : : : : : : :								
Dd	341	CYCCKTCTCPNOPLNPGKCAC-ECATESPOCKLTGKRKFHHQTSCS-----YRRPCTNRBK	393							
Oy	130	ACAACHCRPFKE	140							
		: : : :								
Dd	394	ACEBGFYSSEE	404							

```

14 RESULT
15 US-08-510-133A-35
16 ; Sequence 35, Application US/08510133A
17 ; Patent No. 6221839
18 ; GENERAL INFORMATION:
19 APPLICANT: Altitalo, Kari
20 Joukov, Vladomir
21 TITLE OF INVENTION: Receptor Ligand
22 NUMBER OF SEQUENCES: 35
23 CORRESPONDENCE ADDRESS:
24 ADDRESSEE: Marshali, O'Toole, Gerstein, Murray & Borun
25 STREET: 6300 Sears Tower, 233 South Wacker Drive
26 CITY: Chicago
27 STATE: Illinois
28 COUNTRY: United States of America
29 ZIP: 60606-6402
30 COMPUTER READABLE FORM:
31 MEDIUM TYPE: floppy disk
32 COMPUTER: IBM PC compatible
33 OPERATING SYSTEM: PC-DOS/MS-DOS
34 SOFTWARE: Patentn Release #1.0, Version #1.25
35 CURRENT APPLICATION DATA:
36 APPLICATION NUMBER: US/08/510.133A
37 FILING DATE: 01 Aug-1995
38 CLASSIFICATION: <Unknown>
39 ATTORNEY/AGENT INFORMATION:
40 NAME: Gass, David A.
41 REGISTRATION NUMBER: 38,153
42 REFERENCE/DOCKET NUMBER: 28113/32863
43 TELECOMMUNICATION INFORMATION:
44 TELEPHONE: 312/474-6300
45 TELEFAX: 312/474-0448
46 TELEX: 25-3856
47 INFORMATION FOR SEQ ID NO: 35:
48 SEQUENCE CHARACTERISTICS:
49 LENGTH: 419 amino acids
50 TYPE: amino acid
51 TOPOLOGY: linear
52 MOLECULE TYPE: protein
53 SEQUENCE DESCRIPTION: SEQ ID NO: 35:
54 US-08-510-133A-35

```

```
Query Match          27.3%; Score 244.5; DB 4; Length 419;  
Best Local Similarity 29.8%; Pred. No. 1.2e-15;  
Matches      57; Conservative    26; Mismatches   47; Indels     61; Gaps       8;
```

```
QY        1 IIRRSI-QIPEDRCSSKKLCPIDMLMSDKCKVCLOE-----ENPLACTD-- 47  
||||| :||: |:||::|::::|:| |||  
Db      224 IIRSLPPTLPQ---CQAANKTCPTNIWNNHHICLCIAQEDFMFSSDAGDDSTDSFHIC 280  
         |||  
QY      48 -HSHLOE-----PALGPH-----IMPEDRCE 69  
         ::| |||||  
Db     281 GPNNELDEETQCACVCRAGLRPSGCSPHELDNRSCQCVCKNKLFPSGCANREPEINTCQ 340
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:48:44 ; Search time 78.16 Seconds
(without alignments)
149.113 Million cell updates/sec

Title: US-09-427-657-4_COPY_202_354

Perfect score: 897
Sequence: 1 IIRRSIOIPEDRCRSHSKL.....HCRPEKRAAQPGRSRNP 153

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	244.5	27.3	419	2 S69207	vascular endothelial
2	150.5	16.8	1700	2 S08167	Baldian1 ring 3 pr
3	120.5	13.4	160	2 JQ0542	185K secretory pro
4	116.5	13.0	1187	2 T18355	hypothetical prote
5	108	12.0	748	2 S66129	disintegrin (EC 3.
6	105	11.7	2406	2 A54148	odx protein - frui
7	105	11.7	2515	2 S47008	tenascin-like prote
8	103	11.5	2946	2 T15840	hypothetical prote
9	100	11.1	1746	1 S19694	tenascin precursor
10	99.5	11.1	2437	2 S42612	transmembrane prot
11	98.5	11.0	832	2 A31246	neurogenic protein
12	98.5	11.0	880	2 S00670	neurogenic repetit
13	98.5	11.0	5376	2 T42215	zonadhesin - mouse
14	98	10.9	356	2 A25918	chromodomain - b
15	97.5	10.9	1599	2 T16210	hypothetical prote
16	96.5	10.8	336	2 D69074	polyferredoxin 4x2
17	96.5	10.8	349	2 S57453	integrin beta-3 su
18	96.5	10.8	788	2 T15130	integrin beta-3 su
19	96.5	10.8	2910	2 T42214	otogelin - mouse
20	96	10.7	1106	2 T44598	hypothetical prote
21	95.5	10.6	2629	2 T32735	telomerase-associa
22	95	10.6	1574	2 T13954	MEGF6 protein - ra
23	94.5	10.5	259	2 T21011	hypothetical prote
24	94.5	10.5	1548	2 S34583	serine protease
25	94	10.5	497	2 T27827	hypothetical prote
26	94	10.5	1321	2 JEO352	mucin MUC5B, trach
27	93.5	10.4	833	2 S19087	gene Delta protein
28	93	10.4	253	2 T25768	hypothetical prote
29	93	10.4	2019	1 JQ1322	tenascin precursor

30	93	10.4	2201	2 A32160	tenascin-C - human
31	92.5	10.3	577	2 A60501	thrombomodulin pre
32	92.5	10.3	846	2 A30889	integrin beta chain
33	92.5	10.3	1077	2 T41146	probable cysteine-
34	92.5	10.3	1680	2 A43434	furin (EC 3.4.21.7
35	92	10.3	1113	2 JEO315	low-density lipopr
36	92	10.3	1188	2 D86236	protein F14N23.5 f
37	91.5	10.2	527	2 A42032	epidermal growth f
38	91.5	10.2	572	2 T29880	hypothetical prote
39	91.5	10.2	722	2 T48324	hypothetical prote
40	91.5	10.2	1223	1 TVCHLV	epidermal growth f
41	91.5	10.2	1299	2 T43251	furin (EC 3.4.21.7
42	91	10.1	197	2 T10081	sperm mitochondria
43	91	10.1	680	2 PN0510	integrin beta-3 ch
44	90.5	10.1	379	2 A59180	mnt inhibitory fac
45	90.5	10.1	1620	2 T27283	hypothetical prote

ALIGNMENTS

RESULT 1
S69207
Vascular endothelial growth factor C precursor - human
N:Alternate names: FLT4 ligand DHM
C:Species: Homo sapiens (man)
C:Date: 27-Apr-1996 #sequence, revision 01-Nov-1996 #text_change 08-Oct-1999
C:Accession: S69207; S61795; S71443; S69208; G02659
R:Joukov, V.; Pajusola, K.; Kaipainen, A.; Chillov, D.; Lahtinen, I.; Kukk, E.; Saksel
EMBO J. 15, 1751, 1996
A:Title: Corrigendum: A novel vascular endothelial growth factor, VEGF-C, is a ligand
A:Reference number: S69207; MUID:96203094
A:Accession: S69207
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-419 <JOU>
A:Cross-references: EMBL:X94216; NID:g1177488; PIDN:CA63907.1; PID:e221096; PID:g118
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1995
A>Note: this is a revision to the translation is shown
R:Joukov, V.; Pajusola, K.; Kaipainen, A.; Chillov, D.; Lahtinen, I.; Kukk, E.; Saksel
EMBO J. 15, 290-298, 1996
A:Title: A novel vascular endothelial growth factor, VEGF-C, is a ligand for the FLT4
A:Reference number: S61795; MUID:96178224
A:Accession: S61795
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 70-419 <JOU1>
A>Note: this sequence has been revised in reference S69207
A:Accession: S71443
A:Molecule type: protein
A:Residues: 'X', 104-120 <JOU2>
R:Lee, J.; Gray, A.; Yuan, J.; Luoh, S.M.; Avraham, H.; Wood, W.I.
Submitted to the EMBL Data Library, December 1995
A:Description: Vascular endothelial growth factor related protein (VRP): A ligand and
A:Reference number: S69208
A:Accession: S69208
A:Molecule type: mRNA
A:Residues: 1-419 <LEE>
A:Cross-references: EMBL:U43142; NID:g1150988; PIDN:AAA85214.1; PID:g1150989
R:Morris, J.C.
Submitted to the EMBL Data Library, May 1996
A:Reference number: H01557
A:Accession: G02659
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-419 <MOR>
A:Cross-references: EMBL:U58111; NID:g1373426; PIDN:AA802909.1; PID:g1373427
C:Genetics:
A:Gene: GDB:VEGFC; VRP
A:Cross-references: GDB:3890883; OMIM:601528
F.1-12/Domain: signal sequence #status predicted <SIG>
F.13-102/Domain: propeptide #status predicted <PRO>

A:Residues: 1-748 <GLY>
A:Cross-references: EMBL:Z21961; NID:g1044810; PIDN:CAA79973.1; PID:g1044811
A:Note: this is a revision to the sequence from reference S32205
R:Glynn, P.; Howard, L.
A:Description: Identification of a mammalian member of the metalloproteinase/disintegrin
A:Reference number: S32205
A:Accession: S32205
A:Molecule type: mRNA
A:Residues: 1-113, 'LAM', 117, 'LLLMEDLKDSFRLMVAR', 135-171, 'R', 173-652, 'L' <GLM>
A:Cross-references: EMBL:Z21961
A:Note: this sequence has been revised in reference S66129
C:Superfamily: disintegrin homology
C:Keywords: hydrolase; metalloproteinase; zinc
F:1-13/Domain: signal sequence #status predicted <SIG>
F:14-748/Product: metalloproteinase #status predicted <KMT>
F:456-546/Domain: disintegrin homology <DIS>
F:383,387,393/Binding site: zinc, catalytic (His) #status predicted
F:384/Active site: Glu #status predicted

Query Match 12.0%; Score 108; DB 2; Length 748;
Best Local Similarity 22.3%; Pred. No. 0.13;
Matches 37; Conservative 18; Mismatches 53; Indels 58; Gaps 8;

OY 12 DRCSHSK-KLCP---IDMLDSNCKCYLOENPLAGTEDHSHLOEPALCGPHMFEDDR 67

DB 425 DKLNNKSLSCIRINSOYLEKRNKCYVESGP-----ICGNGWEQGRE 470

OY 68 CBC---VCKTPPCPKDLQHPKNCSC---PECKESETCCOKHKLPHDPDSCSD--- 115

DB 471 CCGGSDCKDCBCYDANQ---PEGKKCKLPGKQCSGSPGCPCTAHCAFKSKTEKCRDSD 529

OY 116 -----RCPFHTRPCASGKTA---CAHK 134

DB 530 CAKEGICNGITALCPASDPKPNFTDCNRTQYICNGCAGSICEKH 575

RESULT 6
A54148
cod protein - fruit fly (Drosophila sp.)

C:Species: Drosophila sp.
C:Date: 02-Aug-1994 #sequence_revision 02-Aug-1994 #text_change 11-Jan-2000
C:Accession: A54148

R:Levine, A.; Bashan-Ahrend, A.; Budai-Hadrian, O.; Gartenberg, D.; Menasheerov, S.; Wild

A:Title: odd Oz: a novel Drosophila pair rule gene.

A:Reference number: A54148; MUID:94243925

A:Accession: A54148

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-2406 <LEV>

C:Genetics:

A:Gene: FlyBase:Ten-m

A:Cross-references: FlyBase:FBgn0004449

C:Superfamily: unassigned EGF-related proteins; EGF homology

F:514-540/Domain: EGF homology <EGF>

F:610-637/Domain: EGF homology <EGF1>

Query Match 11.7%; Score 105; DB 2; Length 2406;
Best Local Similarity 26.2%; Pred. No. 0.54;
Matches 38; Conservative 9; Mismatches 66; Indels 32; Gaps 6;

DB 714 RCNEHQ-CXNGTCLCVTGMNGKHC 737

RESULT 7
S47008
tenascin-like protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 21-Jul-2000
C:Accession: S47008

R:Baumgartner, S.; Martin, D.; Hagios, C.; Chiquet-Ehrismann, R.

EMBO J. 13, 3728-3740, 1994

A:Title: Ten(m), a Drosophila gene related to tenascin, is a new pair-rule gene.

A:Reference number: S47008; MUID:94349920

A:Accession: S47008

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-2515 <BAU>

A:Cross-references: EMBL:X73154; NID:g510505; PIDN:CAA51678.1; PID:g510506

C:Genetics:

A:Gene: FlyBase:Ten-m

A:Cross-references: FlyBase:FBgn0004449

C:Superfamily: unassigned EGF-related proteins; EGF homology

F:298-324/Domain: EGF homology <EGF>

F:394-421/Domain: EGF homology <EGF1>

Query Match 11.7%; Score 105; DB 2; Length 2515;
Best Local Similarity 26.4%; Pred. No. 0.56;
Matches 38; Conservative 9; Mismatches 66; Indels 32; Gaps 6;

OY 10 EEDRCSHSK-KLCPIDMLDSNCKCYLOENPLAGTEDHSHLOEPALCGPHMFEDDR 69

DB 390 EYVDCPHPN--CSGHGFCADGTCICKKWKGPDCATMDALQCLPDCSGHGFDDYDT 447

OY 70 CYCKTPPCPKDLQHPKNCSCFECKESETCCOKHKLPHDPDSCSD----- 115

DB 448 CTCBAKMSGD-----DCS---KELCDLDCGQHGRCESDADACDPDEWGEYCNTRLCDV 497

OY 116 RCPFHTRPCASGKTA---AKHC 135

DB 498 RCNEHQ-CXNGTCLCVTGMNGKHC 521

RESULT 8
T15840
hypothetical protein C54G7.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T15840

R:Du Z.

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2946 <DUZ>

A:Cross-references: EMBL:U04010; NID:g1065453; PID:g1065455; PIDN:AAA81392.1; CESP:CS

A:Gene: CESP:C54G7.3

A:Introns: 16/1; 53/2; 92/1; 160/3; 295/1; 346/1; 392/1; 440/1; 475/3; 579/1; 615/3;

46/1; 2561/2; 2603/2; 2626/2; 2665/2; 2716/3; 2804/3; 2884/3

Query Match 11.5%; Score 103; DB 2; Length 2946;
Best Local Similarity 26.4%; Pred. No. 0.91;
Matches 38; Conservative 15; Mismatches 57; Indels 34; Gaps 8;

OY 3 RSTIQPEDRCSHSK-KLCPIDMLDSNCKCYLOENPLAGTEDHSHLOEPALC----- 57

DB 1068 RPTSTIAGHCTHNS-CPSEFSFCFSNCNC-----MAGFRATSGICEPAIAVAGPC 1119

500670
 A:Cross-references: EMBL:X06289; NID:g7852; PID:g7853
 A:Residues: 1-880 <VAE>
 A:Molecule type: mRNA
 A:Accession: S00670
 A:Title: The neurogenic Delta of *Drosophila melanogaster* is expressed in neurogenid
 A:Reference number: S00670
 A:Accession: S00670
 A:Molecule type: mRNA
 A:Residues: 1-880 <VAE>
 A:Cross-references: EMBL:X06289; NID:g7852; PID:g7853
 R:Knust, E.; Dietrich, U.; Teppas, U.; Bremer, K.A.; Weigel, D.; Vaessin, H.; Campos-Ort
 EMO J. 6, 761-766, 1987
 A:Title: EGF homologous sequences encoded in the genome of *Drosophila melanogaster*, and
 A:Reference number: A91081; MUID:87218537
 A:Accession: A26637
 A:Molecule type: mRNA
 A:Residues: 422-436, 'ET', 'A', 439-458, 'A', 460-489, 'T', 491-521 <KN>
 A:Cross-references: GB:X05140; NID:g7851; PIDN:CA8786.1; PID:g929563
 C:Genetics:
 A:Gene: Delta; DL
 A:Cross-references: FlyBase:FBgn0000463
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 C:Keywords: transmembrane protein
 F:1-18/Domain: signal sequence #stratus predicted <SIG>
 F:19-88/Product: neurogenic repetitive locus delta protein #status predicted <MAT>
 F:457-468/Domain: EGF homology <EGF1>
 F:533-564/Domain: EGF homology <EGF2>

[illegible]

	Query Match	Similarity	11.0%;	Score 98.5;	DB 2,	Length 5376;
	Best Local	Similarity	25.4%;	Pred No. 3.3;		
	Matches	43;	Conservative	15;	Mismatches	66;
					Indels	43;
					Gaps	9;
Qy	4	RSIQIPEDRCSHSK-----LCPIDMLMNSKCKVLOENP---LAGT--EDH	48			
Db	2673	RFQCPGSGHYCKDIKDDASNCTETILQCPDHSILYTHCLPSCILCSDDPGLCRGTSPEAP	2732			
Qy	49	SHLQRPALCGPMMDEDEDC-----ECVCK-----TPCKKDIQH--PANCSC-----	89			
Db	2733	STCKGCGVCDPPYVLNSKCKVLRIKCGCKDAAGVLIIPAKTMINRGSCSCGCMGAIQC	2792			
Qy	90	--FECKESLLETCCOKHKLFHPDTCSCEDRCPEHTPRCAGSKAPCAKHKR	136			
Db	2793	QNEKPS--EATYQDMEDGNSNCTSLPDCRPH-----SHYINCLPTQC	2834			

RESULT 14
A25918
thrombomodulin - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 30-Jun-1988 #sequence-revision 30-Jun-1988 #text-change 16-Jul-1999
C:Accession: A25918
R:Jackman, R.W.; Beeler, D.L.; Vandewater, L.; Rosenberg, R.D.
Proc. Natl. Acad. Sci. U.S.A. 83, 8834-8838, 1986
A:Title: Characterization of a thrombomodulin cDNA reveals structural similarity to t
A:Reference number: A25918; MUID:87067408
A:Accession: A25918
A:Molecule type: mRNA
A:Residues: 1-356 <JLAC>
A:Cross-references: GB:A4657; NID:q163762; PIDN:AAA30785.1; PID:q163763
C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology
C:Keywords: transmembrane protein
F:21-56/Domain: EGF homology <EG1>
F:64-97/Domain: EGF homology <EG2>
F:103-136/Domain: EGF homology <EG3>
F:143-178/Domain: EGF homology <EG4>
F:182-213/Domain: EGF homology <EG5>
F:219-253/Domain: EGF homology <EG6>

[illegible]

RESULT 15
T16210
hypothetical protein F30H5_3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16210
R:Pauley, A.; Steiljes, L.
Submitted to the EMBL Data Library, June 1995
:Description: The sequence of *C. elegans* cosmid F30H5.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:51:49 ; Search time 42.92 Seconds
(without alignments)
122.113 Million cell updates/sec

Title: US-09-427-657-4_COPY_202_354
Perfect score: 897
Sequence: 1 IIRRSIQIPEDRCSHKKL.....HCRFPKRAQGPMSRKNP 153

Scoring table: BLAST62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	244.5	27.3	419	1	VEGC_HUMAN
2	236.5	26.4	415	1	VEGC_MOUSE
3	150.5	16.8	1700	1	BAR3_CHITE
4	105	11.7	191	1	WAP_MACEU
5	100	11.1	1746	1	TENA_PIG
6	99.5	11.1	2437	1	NOTC_BRARE
7	98	10.9	356	1	TRBM_BOVIN
8	96.5	10.8	1696	1	PKC5_BRACL
9	94.5	10.5	1877	1	PKC5_MOUSE
10	93.5	10.4	833	1	DL_DROME
11	93	10.4	2201	1	TENA_HUMAN
12	92.5	10.3	577	1	TRBM_MOUSE
13	92.5	10.3	846	1	ITBX_DROME
14	92.5	10.3	1680	1	FUR2_DROME
15	92	10.3	703	1	SELFB_BOVIN
16	91.5	10.2	402	1	EGFR_CHICK
17	91.5	10.2	714	1	DL11_RAT
18	91.5	10.2	722	1	DL11_MOUSE
19	91	10.1	197	1	MCS_MOUSE
20	91	10.1	787	1	WIF1_MOUSE
21	90.5	10.1	379	1	WIF1_HUMAN
22	90	10.0	1607	1	MIPR_LYMT
23	89.5	10.0	913	1	PKC5_HUMAN
24	89.5	10.0	1426	1	EGFR_DROME
25	89	9.9	571	1	DISJ_BOVIA
26	89	9.9	1394	1	TGFB_HUMAN
27	89	9.9	2871	1	FBN1_BOVIN
28	89	9.9	2871	1	FBN1_HUMAN
29	89	9.9	3051	1	FNX3_CAEEL
30	88.5	9.9	1257	1	ERR2_RAT
31	88.5	9.9	2414	1	P300_HUMAN
32	88	9.8	381	1	SELFB_HUMAN
33	88	9.8	575	1	TRBM_HUMAN

RESULT	ID	VEGC_HUMAN	STANDARD	PRT	419 AA
AC	P49767				
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	01-OCT-1996	(Rel. 40, Last annotation update)			
DE	VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR (VEGF-C) (VASCULAR				
DE	ENDOTHELIAL GROWTH FACTOR RELATED PROTEIN) (VRF) (FLT4 LIGAND) (FLT4-				
DE	L).				
GN	VEGFC.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
OX	NCBI_TaxID=9606;				
RM	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 103-120.				
RX	MEDLINE=96176224; PubMed=8617204;				
RA	Joukov V., Pajusola K., Kaipainen A., Chillov D., Lahtinen I., Kukk E.,				
RA	Saksela O., Kalkkinen N., Alitalo K.;				
RT	"A novel vascular endothelial growth factor, VEGF-C, is a ligand for				
RT	the Flt4 (VEGFR-3) and KDR (VEGFR-2) receptor tyrosine kinases.";				
RL	EMBO J. 15:290-298(1996).				
RN	[2]				
RP	ERRATUM.				
RX	MEDLINE=96203094; PubMed=8612600;				
RA	Joukov V., Pajusola K., Kaipainen A., Chillov D., Lahtinen I., Kukk E.,				
RA	Saksela O., Kalkkinen N., Alitalo K.;				
RL	EMBO J. 15:1751-1751(1996).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=96312526; PubMed=8700872;				
RA	Lee J., Gray A., Yuan J., Luo S.-M., Avraham H., Wood W.I.;				
RT	"Vascular endothelial growth factor-related protein: a ligand and				
RT	specific activator of the tyrosine kinase receptor Flt4.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 93:1988-1992(1996).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	Fitz L., Morris J.C., Towler P.S., Long A.J., Greco R.,				
RA	Burgess P., Giannotti J., Claretta A., Hennessey D., Kovacic S.,				
RA	Fitzgerald M., Scaltreto H., Welch N., Neben S., Finerly H.,				
RA	Zollner R., Wang J., Nickbarg E., Gassaway R., Turner K.,				
RL	Wood C.R.;				
RT	Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.				
CC	-!- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL				
CC	CELL GROWTH.				
CC	-!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.				
CC	-!- PTM: PROBABLY PROTEOLYTICALLY PROCESSED IN THE C-TERMINUS.				
CC	-!- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.				
CC	-!- This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				

or send an email to license@isb-sib.ch).

CC EMBL: X94216; CAA63907.1; -
 CC EMBL: U43142; AAB85214.1; -
 DR EMBL: U58111; AAB02909.1; -
 DR HSSP: P15692; 1VPF.
 DR MIM: 601528; -
 DR InterPro: IPR000072; -
 DR InterPro: IPR002400; -
 DR Pfam: PF00341; PDGF_1.
 DR PRINTS: PR00438; GFCSKNOT.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS50278; PDGF_2; 1.
 DR Mitogen; Growth factor; Glycoprotein; Signal; Repeat.
 KW SIGNAL
 FT PROPEP 1 102 POTENTIAL.
 FT CHAIN 103 419 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
 FT DOMAIN 275 365 4 X 24 AA TANDEM REPEATS.
 FT REPEAT 275 298 1.
 FT REPEAT 299 322 2.
 FT REPEAT 323 346 3.
 FT REPEAT 347 365 4 (PARTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 419 AA; 46883 MW; 9F598719DB3E014F CRC64;

Query Match 27.3%; Score 244.5; DB 1; Length 419;
 Best Local Similarity 29.8%; Pred. No. 1.6e-14;
 Matches 57; Conservative 26; Mismatches 47; Indels 61; Gaps 8;

OY 1 IIRRSI--QIPEEDRCSHKKLPIDMLDSNKKCVLOE-----ENPLAGTED-- 47
 DB 224 IIRSLPTLPQ---CQANKTCPTNYMNMNHCICLAQEDMFSSDAGDSDTFPHDLC 280
 OY 48 --HSHLOE-----PALCGPH-----MPEDEDC 69
 DB 281 GPNKLEDETCQCVKAGLRPASCGRHKELDNRSCQCVCKNKLFPSCGANREPEENTCQ 340
 OY 70 CVCKTPCKDLIQHPKNSCFECKESLETCCQKHKLPHDPDSCDCRCPFHRPCASGKT 129
 DB 341 CVCKTCCPRNPLNGKAC--ECTESPCKLKGKFFHQTCSC-----YRRPCTNRK 393
 OY 130 ACAKHCRPEKE 140
 DB 394 ACEPGFSYSEE 404

RESULT 2
 VEGF_MOUSE STANDARD; PRT; 415 AA.
 AC P97953;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR (VEGF-C) (FLT4 LIGAND) (FLT4-L).
 GN VEGFC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BAIB/C;
 RX MEDLINE=97164697; PubMed=9012504;
 RA Kukk E., Lymbousaki A., Taira S., Kaipainen A., Jeltsch M., Joukov V., Alltalo K.;
 RT "VEGF-C receptor binding and pattern of expression with VEGFR-3 suggests a role in lymphatic vascular development."; Development 122:3829-3837(1996).
 RL [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=BAIB/C;
 RX MEDLINE=97388482; PubMed=9247316;
 RA Fitz L.J., Morris J.C., Towler P., Long A., Burgess P., Greco R., Wang J., Gassaway R., Nickbarg E., Kovacic S., Chiarlitta A., Giannotti J., Flinerty H., Zollner R., Beier D.R., Leak L.V., Turner K.J., Wood C.R.;
 RA "Characterization of murine flt4 ligand/VEGF-C"; Oncogene 15:613-618(1997).
 RL [1]
 CC -1- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL CELL GROWTH.
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----

CC EMBL: U73620; AAC52984.1; -
 CC EMBL: U58112; AAB46707.1; -
 DR HSSP: P15692; 1VPF.
 DR MGD: MGI:109124; Vegfc.
 DR InterPro: IPR000072; -
 DR InterPro: IPR002400; -
 DR Pfam: PF00341; PDGF_1.
 DR PRINTS: PR00438; GFCSKNOT.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS50278; PDGF_2; 1.
 DR Mitogen; Growth factor; Glycoprotein; Signal; Repeat.
 KW SIGNAL
 FT PROPEP 1 98 POTENTIAL.
 FT CHAIN 99 415 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
 FT DOMAIN 271 361 4 X 24 AA TANDEM REPEATS.
 FT REPEAT 271 294 1.
 FT REPEAT 295 318 2.
 FT REPEAT 319 342 3.
 FT REPEAT 343 361 4 (PARTIAL).
 FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 415 AA; 46471 MW; D9D3DD3CEC659D6 CRC64;

Query Match 26.4%; Score 236.5; DB 1; Length 415;
 Best Local Similarity 30.1%; Pred. No. 7.8e-14;
 Matches 56; Conservative 24; Mismatches 41; Indels 65; Gaps 9;

OY 1 IIRRSI--QIPEEDRCSHKKLPIDMLDSNKKCVLOE-----ENPLAGTED-- 47
 DB 220 IIRSLPTLPQ---CQANKTCPTNYMNMNHCICLAQEDMFSSDAGDSDTFPHDLC 276
 OY 48 --HSHLOE-----PALCGPH-----MPEDEDC 69
 DB 277 GPNKLEDETCQCVKAGLRPASCGRHKELDNRSCQCVCKNKLFPSCGANREPEENTCQ 336
 OY 70 CVCKTPCKDLIQHPKNSCFECKESLETCCQKHKLPHDPDSCDCRCPFHRPCASGKT 129
 DB 337 CVCKTCCPRNPLNGKAC--ECTENTCKFLKGKFFHQTCSC-----YRRPCANR-- 387
 OY 130 ACAKHCR 135
 DB 388 --LRHC 391

RESULT 3
 BAR3_CHITE STANDARD; PRT; 1700 AA.
 ID BAR3_CHITE
 AC Q03376;
 DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE BALBIANI RING PROTEIN 3 PRECURSOR.
 GN BR3
 OS Chironomus tentans (Midge).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
 OC Chironomidae; Chironominae; Chironomus.
 OX NCBI_TaxId=7153;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=90172404; PubMed=1689777;
 RA Paulsson G., Lendahl U., Galli J., Ericsson C., Mieslander L.;
 RT "The Balbiani ring 3 gene in Chironomus tentans has a diverged
 repetitive structure split by many introns.";
 RL J. Mol. Biol. 211:331-349(1990).
 CC -1- FUNCTION: USED BY THE LARVAE TO CONSTRUCT A SUPRAMOLECULAR
 CC STRUCTURE, THE LARVAL TUBE. BALBIANI RING PROTEIN 3 COULD PLAY A
 CC ROLE AS A TRANSPORT PROTEIN THAT BINDS TO OTHER PROTEINS
 CC INTRACELLULARLY AND IN THE GLAND LUMEN IN ORDER TO PREVENT THESE
 CC FROM FORMING WATER-INSOLUBLE FIBERS TOO EARLY.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- TISSUE SPECIFICITY: SALIVARY GLAND.
 CC -1- DOMAIN: HAS 82 APPROXIMATE REPEATS OF CYS-X-CYS-X-CYS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X52263; CA36506.1; -
 DR PIR: S08167; S08167.
 DR HSSP: P18055; 2MRB.
 DR InterPro: IPR000853; -
 DR PRINTS: PR00876; MTNEMATODE.
 KM Repeat; Signal.
 FT SIGNAL 1
 FT CHAIN 1 ? POTENTIAL.
 FT SEQUENCE 1700 AA; 186145 MW; 3420282521B0815 CRC64;
 SO
 Query Match 16.8%; Score 150.5; DB 1; Length 1700;
 Best Local Similarity 26.2%; Pred. No. 6.9e-06;
 Matches 42; Conservative 22; Mismatches 55; Indels 41; Gaps 8;
 QY 13 RCHSHKTLCPIDM-----LWDSNKKCYLQENPLAG-----TEHSH 49
 DB 1246 RC-----VCPKMEKPADNCKTKMWNDEMCKVCCKPGCGEGCKGVKMNANTSCCECPA 1300
 QY 50 HLOEPALCGPHMFDEDRCEVCVCTP-----CPKDLIHPKNCSEFECKSELETCCOKHK 104
 DB 1301 DKAKPASGCKSMKWDSCCKCKMPGCGCPNQNOMMEKNC---ECKSATGNCPPAQ 1357
 QY 105 LEHPDTSCEDRCPPHTRPCASGKTACAKHCR--FPKEKR 142
 DB 1358 TWSOTCOCS--CP-ATGKCTGAQVWCCKACKCVCPAQKK 1394
 RESULT 4
 WAP_MACEU STANDARD; PRT; 191 AA.
 ID WAP_MACEU
 AC Q9NOL8;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE WHEY ACIDIC PROTEIN PRECURSOR (TWAP).
 GN WAP.
 OS Macropus eugenii (Tamar wallaby).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
 OX NCBI_TaxId=9315;
 RN [1]
 RP SEQUENCE FROM N.A., AND 3D-STRUCTURE MODELLING.
 RC TISSUE=Lactating mammary gland;
 RX MEDLINE=20390063; PubMed=10801834;
 RA Simpson K.J., Ranganathan S., Fisher J.A., Janssens P.A., Shaw D.C.,
 RA Nicholas K.R.;
 RT "The gene for a novel member of the whey acidic protein family encodes
 three four-disulfide core domains and is asynchronously expressed
 during lactation.";
 RL J. Biol. Chem. 275:23074-23081(2000).
 CC -1- FUNCTION: COULD BE A PROTEASE INHIBITOR.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- TISSUE SPECIFICITY: MILK-SPECIFIC. MAJOR PROTEIN COMPONENT OF MILK
 CC WHEY.
 CC -1- SIMILARITY: BELONGS TO THE WAP-TYPE 'FOUR-DISULFIDE CORE' FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AJ005356; CAB90357.1; -
 DR PDB: 1TWP; 13-JAN-00.
 DR PROSITE: PS00317; 4_DISULFIDE_CORE; 3.
 KM Milk; Whey; Protease inhibitor; Repeat; Signal; 3D-structure.
 FT SIGNAL 1
 FT CHAIN 1 191
 FT DOMAIN 28 191 WHEY ACIDIC PROTEIN.
 FT DOMAIN 77 125
 FT DOMAIN 132 174 WAP 2.
 FT DISULFID 29 59 BY SIMILARITY.
 FT DISULFID 42 63 BY SIMILARITY.
 FT DISULFID 46 58 BY SIMILARITY.
 FT DISULFID 52 67 BY SIMILARITY.
 FT DISULFID 78 114 BY SIMILARITY.
 FT DISULFID 97 118 BY SIMILARITY.
 FT DISULFID 101 113 BY SIMILARITY.
 FT DISULFID 107 122 BY SIMILARITY.
 FT DISULFID 133 163 BY SIMILARITY.
 FT DISULFID 140 167 BY SIMILARITY.
 FT DISULFID 150 162 BY SIMILARITY.
 FT DISULFID 156 171 BY SIMILARITY.
 FT CARBOHYD 69 69 N-LINKED (GLCNAC...)(POTENTIAL).
 SO SEQUENCE 191 AA; 21133 MW; BF908B1ADABBE50D CRC64;
 Query Match 11.7%; Score 105; DB 1; Length 191;
 Best Local Similarity 24.7%; Pred. No. 0.01;
 Matches 38; Conservative 22; Mismatches 58; Indels 36; Gaps 10;
 QY 10 EEDRSHSKTLCPIDMIMNSNKKCYLQEN---PLAGTEHDS--HLOEPALCGPHMMD 64
 DB 48 DDASCPQNNRCQRCQCSW---LCMNTQKDGCLCPATVSHSSSEQQRKQLCDKTCTKD 104
 QY 65 ---EDRCEVCVCTPCKDLIHPKNCSEF-----ECKSELETCCOKHKLFHPDTSC 113
 DB 105 LGCEGKAKC-CASSCGQ-----TCFMPYKAKPGRCRPATVGLCPCEKKSFMH--TCOR 152
 QY 114 EDRCPHTRPCASGKTACAKHCR--FPKEKRAQ 145
 DB 153 DDQCKENKKCCSS---ACGRCTNPFPEYEYEAQ 183
 RESULT 5
 TENA_PIG STANDARD; PRT; 1746 AA.
 ID TENA_PIG
 AC Q29116; P98142;
 DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (rel. 38, last sequence update)
 DT 01-OCT-2000 (rel. 40, last annotation update)
 DE TENASCIN PRECURSOR (TN) (HEXABRACHION) (CYTONECTIN) (NEURONECTIN)
 DE (GME) (J1) (MOTOFENDINUS ANTIGEN) (GLIOMA-ASSOCIATED-EXTRACELLULAR
 DE MATRIX ANTIGEN) (GP 150-225) (TENASCIN-C) (TN-C) (P230).
 GN HXB.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Submaxillary gland;
 RX MEDLINE=92104189; PubMed=1722152;
 RA Nishii T., Weinstein J., Gillespie W.M., Paulson J.C.;
 RT "Complete primary structure of porcine tenascin: detection of
 RT tenascin transcript in adult submaxillary glands.";
 RL Eur. J. Biochem. 202:643-646(1991).
 RN [2]
 RP SEQUENCE OF 813-825; 887-917; 998-1011; 1597-1608 AND 1719-1730.
 RC TISSUE=Fetal brain;
 RX MEDLINE=98158323; PubMed=9498558;
 RA Wakatsuki S., Ho S.H., Arioka M., Yamasaki M., Kitamoto K.;
 RT "Isolation and characterization of a 230 kDa protein (p230)
 RT specifically expressed in fetal brains: its involvement in neurite
 RT outgrowth from rat cerebral cortex neurons grown on monolayer of
 RT astrocytes.";
 RL J. Biochem. 122:1146-1152(1997).
 CC -1- FUNCTION: SAM (SUBSTRATE-ADHESION MOLECULE) THAT APPEARS TO
 CC INHIBIT CELL MIGRATION. MAY PLAY A ROLE IN SUPPORTING THE GROWTH
 CC OF EPITHELIAL TUMORS.
 CC -1- FUNCTION: PLAYS A ROLE DURING EARLY BRAIN DEVELOPMENT PARTICULARLY
 CC IN GROWTH CONE GUIDANCE. INVOLVED IN NEURITE OUTGROWTH FROM
 CC CORTICAL NEURONS GROWN ON THE MONOLAYER OF ASTROCYTES.
 CC -1- SUBUNIT: HEXAMERIC. A HOMOTRIMER MAY BE FORMED IN THE TRIPLE
 CC COILED-COIL REGION AND MAY BE STABILIZED BY DISULFIDE RINGS AT
 CC BOTH ENDS. TWO OR SUCH HALF-HEXABRACHIONS MAY BE DISULFIDE LINKED
 CC WITHIN THE CENTRAL GLOBULE.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; MAJOR, MINOR-1 AND MINOR-2
 CC (SHOWN HERE) ARE PRODUCED BY ALTERNATIVE SPLICING. THEY ARE
 CC PRODUCED IN A TISSUE- AND TIME-SPECIFIC MANNER DURING DEVELOPMENT.
 CC -1- TISSUE SPECIFICITY: SUBMAXILLARY GLANDS AND BRAIN.
 CC -1- DEVELOPMENTAL STAGE: PREDOMINANTLY EXPRESSED IN THE EMBRYONIC AND
 CC EARLY POSTNATAL STAGES. LITTLE OR NO DETECTION IN ADULT BRAIN.
 CC -1- INDUCTION: BY TGF-BETA.
 CC -1- SIMILARITY: CONTAINS 15-EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 11 FIBRONECTIN TYPE-III-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 FIBRINOGEN-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X61599; CAA43796.1; -
 DR InterPro: IPR000561; -
 DR InterPro: IPR001777; -
 DR InterPro: IPR002049; -
 DR InterPro: IPR002181; -
 DR Pfam: PF000008; EGF_14;
 DR Pfam: PF00147; fibrinogen_C_1;
 DR Pfam: PF00041; fn3_10;
 DR PRINTS: PRO0011; EGF_LAMININ.
 DR PRINTS: PRO0014; FNTYPEPILI.
 DR PROSITE: PS00022; EGF_1; 15.
 DR PROSITE: PS01186; EGF_2; 14.
 KM Glycoprotein; Cell adhesion; Repeat; EGF-like domain; Coiled coil;
 KM Extracellular matrix; Alternative splicing; Signal.
 FT SIGNAL 1 22 POTENTIAL.
 FT

FT CHAIN 23 1746 TENASCIN.
 FT DOMAIN 23 185 INVOLVED IN HEXAMER FORMATION.
 FT DOMAIN 118 145 COILED COIL (POTENTIAL).
 FT DOMAIN 174 186 EGF-LIKE 1 (INCOMPLETE).
 FT DOMAIN 187 217 EGF-LIKE 2.
 FT DOMAIN 218 249 EGF-LIKE 3.
 FT DOMAIN 250 280 EGF-LIKE 4.
 FT DOMAIN 281 311 EGF-LIKE 5.
 FT DOMAIN 312 342 EGF-LIKE 6.
 FT DOMAIN 343 373 EGF-LIKE 7.
 FT DOMAIN 374 404 EGF-LIKE 8.
 FT DOMAIN 405 435 EGF-LIKE 9.
 FT DOMAIN 436 466 EGF-LIKE 10.
 FT DOMAIN 467 497 EGF-LIKE 11.
 FT DOMAIN 498 528 EGF-LIKE 12.
 FT DOMAIN 529 559 EGF-LIKE 13.
 FT DOMAIN 560 589 EGF-LIKE 14.
 FT DOMAIN 590 620 EGF-LIKE 15.
 FT DOMAIN 621 710 FIBRONECTIN TYPE-III 1.
 FT DOMAIN 711 801 FIBRONECTIN TYPE-III 2.
 FT DOMAIN 802 891 FIBRONECTIN TYPE-III 3.
 FT DOMAIN 892 983 FIBRONECTIN TYPE-III 4.
 FT DOMAIN 984 1071 FIBRONECTIN TYPE-III 5.
 FT DOMAIN 1072 1162 FIBRONECTIN TYPE-III 6.
 FT DOMAIN 1163 1253 FIBRONECTIN TYPE-III 7.
 FT DOMAIN 1254 1342 FIBRONECTIN TYPE-III 8.
 FT DOMAIN 1343 1430 FIBRONECTIN TYPE-III 9.
 FT DOMAIN 1431 1518 FIBRONECTIN TYPE-III 10.
 FT DOMAIN 1527 1733 FIBRINOGEN BETA/GAMMA.
 FT DISULFID 64 64 INTERCHAIN (POTENTIAL).
 FT DISULFID 150 200 BY SIMILARITY.
 FT DISULFID 194 205 BY SIMILARITY.
 FT DISULFID 207 216 BY SIMILARITY.
 FT DISULFID 221 231 BY SIMILARITY.
 FT DISULFID 225 236 BY SIMILARITY.
 FT DISULFID 238 247 BY SIMILARITY.
 FT DISULFID 252 263 BY SIMILARITY.
 FT DISULFID 256 268 BY SIMILARITY.
 FT DISULFID 270 279 BY SIMILARITY.
 FT DISULFID 284 294 BY SIMILARITY.
 FT DISULFID 288 299 BY SIMILARITY.
 FT DISULFID 301 310 BY SIMILARITY.
 FT DISULFID 315 325 BY SIMILARITY.
 FT DISULFID 319 330 BY SIMILARITY.
 FT DISULFID 332 341 BY SIMILARITY.
 FT DISULFID 346 356 BY SIMILARITY.
 FT DISULFID 350 361 BY SIMILARITY.
 FT DISULFID 363 372 BY SIMILARITY.
 FT DISULFID 377 387 BY SIMILARITY.
 FT DISULFID 381 392 BY SIMILARITY.
 FT DISULFID 394 403 BY SIMILARITY.
 FT DISULFID 408 418 BY SIMILARITY.
 FT DISULFID 412 423 BY SIMILARITY.
 FT DISULFID 425 434 BY SIMILARITY.
 FT DISULFID 439 449 BY SIMILARITY.
 FT DISULFID 443 454 BY SIMILARITY.
 FT DISULFID 456 465 BY SIMILARITY.
 FT DISULFID 470 480 BY SIMILARITY.
 FT DISULFID 474 485 BY SIMILARITY.
 FT DISULFID 487 496 BY SIMILARITY.
 FT DISULFID 501 511 BY SIMILARITY.
 FT DISULFID 505 516 BY SIMILARITY.
 FT DISULFID 518 527 BY SIMILARITY.
 FT DISULFID 532 542 BY SIMILARITY.
 FT DISULFID 536 547 BY SIMILARITY.
 FT DISULFID 549 558 BY SIMILARITY.
 FT DISULFID 563 573 BY SIMILARITY.
 FT DISULFID 567 578 BY SIMILARITY.
 FT DISULFID 580 589 BY SIMILARITY.
 FT DISULFID 594 604 BY SIMILARITY.
 FT DISULFID 598 609 BY SIMILARITY.
 FT DISULFID 611 620 BY SIMILARITY.
 FT CARBOHYD 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT

DT	01-OCT-2000 (Rel. 40, last annotation update)
DE	PROTEIN CONVERTASE SUBTILISIN/KEXIN TYPE 5 PRECURSOR (EC 3.4.21.-)
DE	(PROTEIN CONVERTASE PC6-LIKE) (APC6).
GN	PC6.
OS	Brachiostoma californiensis (California lancelet) (Amphioxus).
OC	Eukaryota; Metazoa; Chordata; Cephalochordata; Brachiostomidae;
CC	Brachiostoma.
OX	NCBJ TaxID=7738;
RN	[1]
RP	SEQUENCE FROM N.A. (ISOFORM A; ISOFORM B AND ISOFORM C).
RX	MEDLINE=20175281; PubMed=10708866;
RA	Oliva A.A. Jr., Chan S.J., Steiner D.F.: "
RT	"Evolution of the prohormone convertases: identification of a
RL	homologue of pc6 in the protochordate amphioxus.";
Biochim. Biophys. Acta 1477:338-348(2000)."	
-I-	FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENDOPEPTIDASE ACTIVITY WITHIN THE CONSTITUTIVE AND REGULATED SECRETORY PATHWAY. CAPABLE OF CLEAVAGE AT THE RK/RKRR CONSSENSUS MOTIF (BY SIMILARITY).
-I-	CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR PROPEPTEINS BY CLEAVAGE OF ARG-XAA-YAA-ARG-I'-ZAA BONDS, WHERE XAA CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
-I-	SUBCELLULAR LOCATION: ISOFORM A AND ISOFORM C ARE SECRETED. ISOFORM B IS A TYPE I MEMBRANE PROTEIN.
-I-	ALTERNATIVE PRODUCTS: 3 ISOFORMS: A, B (SHOWN HERE) AND C; ARE PRODUCED BY ALTERNATIVE SPLICING.
-I-	DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE ASSISTING THE FOLDING OF THE ZYMOGEN WITHIN THE ENDOPLASMIC RETICULUM.
-I-	SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE SUBTILASE FAMILY.
-I-	SIMILARITY: CONTAINS 1 HOMO E/P DOMAIN.
-----	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).
CC	EMBL: AF184615; AAF26300.1; -
DR	EMBL: AF184616; AAP26301.1; -
DR	EMBL: AF184617; AAF26302.1; -
DR	PROSITE: PS00136; SUBTILASE_ASP; FALSE_NEG.
DR	PROSITE: PS00137; SUBTILASE_HIS; 1.
DR	PROSITE: PS00138; SUBTILASE_SER; 1.
KW	Hydrolase; serine protease; Glycoprotein; zymogen; signal; cleavage on pair of basic residues; Repeat; Alternative splicing; Transmembrane.
KM	
FT	SIGNAL 1 25 POTENTIAL.
ET	POTENTIAL.
ET	PROPEP 26 110 PROTEIN CONVERTASE SUBTILISIN/KEXIN
ET	CHAIN 111 1696 TYPE 5.
FT	DOMAIN 111 1618 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 1619 1639 POTENTIAL.
FT	DOMAIN 1640 1696 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 111 488 CATALYTIC.
FT	HOMO B.
FT	CYS-RICH MOTIF (CRM) REGION.
FT	SITE 110 111 CLEAVAGE (AUTO-) (BY SIMILARITY).
FT	ACT_SITE 192 192 CHANGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 203 203 CHANGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 407 407 CHANGE RELAY SYSTEM (BY SIMILARITY).
FT	CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 529 529 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 885 885 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPLIC 1259 1323 DDITLDGECTSCSGPEVMDRKKCKCAHPKCECSDEX DDTTCAGNDGFLTDASCSAGCP -> AENOMASCPRA PREVSVAELALGHLRYSLDVPOSMSPDVLGADRARTL TTATSAAGRA (IN ISOFORM C).
FT	MISSING (IN ISOFORM C).
FT	CHPCKCRSDCHREDDTK-> IARCVDRRRCWCMDLVLRNFNC
FT	VARSPPLIC 1324 1696
FT	VARSPPLIC 1288 1343

```

FT FT VARSPLIC 1344 1696 VARYEVKRCCTCKLYMEDRPMRRGSSOPTGGRN (IN
FT FT SEQUENCE 1696 AA: 188410 MW: 281GBE1784257CDB CRC64;
SO SEQUENCE 1696 AA: 188410 MW: 281GBE1784257CDB CRC64;

Query Match 10.8%; Score 96.5; DB 1; Length 1696;
Best Local Similarity 19.9%; Pred. No. 0.33;
Matches 35; Conservative 23; Mismatches 57; Indels 61; Gaps 8;

QY 9 PEEDRSHSKKCLIPIMDMNSKCKVLOENPLAGED-HSLLOEPALCGPMMDEDR 67
DB 988 PADNOCY-----LCEERAPPTDGRGQSGSDTDEACACGCHSCSEGPDTIC-----DS 1035
QY 68 CE-----CYCKTPCPKDLIQHKNCSCEFEKESLETC-----COKHKLHP 108
DB 1036 CDDDYLTETCTCARTNCPSFTYPPDDQDRCHRCNCEACGDPNNONCSCEGEYKTP 1095
QY 109 DTCS-----CEDRC-----PHTRPCASG-----KTACAKHC 135
DB 1096 DGCSTGCPNRYKDYDNKECKPCDSSCGTSCGSPFHLSCADGDFLHSSCRKSTC 1151

RESULT 9
PC5_MOUSE STANDARD; PRT: 1877 AA.
AC 004592; 062040;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROTEIN CONVERTASE SUBTILISIN/KEXIN TYPE 5 PRECURSOR (EC 3.4.21.-)
DE (PROTEIN CONVERTASE PC5) (SUBTILISIN/KEXIN-LIKE PROTEASE PC5)
DE (CONVERTASE PC5) (PC6) (SUBTILISIN-LIKE PROTEIN CONVERTASE 6)
DE (SPC6).
GN PCSK5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE OF 330-1877 FROM N.A. (LONG ISOFORM).
RC STRAIN=ICR; TISSUE=Intestine;
RX MEDLINE-93327934; PubMed-835106;
RA Nakagawa T., Murakami K., Nakayama K.;
RT "Identification of an isoform with an extremely large Cys-rich region
RT of PC6, a Kex2-like processing endoprotease."
RT FEBS Lett. 327:165-171(1993).
RN
[2]
RN SEQUENCE FROM N.A. (SHORT ISOFORM).
RC TISSUE=Brain, and Intestine;
RX MEDLINE-93324489; PubMed-8468318;
RA Nakagawa T., Hosaka M., Torii S., Watanabe T., Murakami K.,
RA Nakayama K.;
RT "Identification and functional expression of a new member of the
RT mammalian Kex2-like processing endoprotease family: its striking
RT structural similarity to PRCE4."
RT J. Biochem. 113:132-135(1993).
RN
[3]
RN SEQUENCE FROM N.A. (SHORT ISOFORM).
RC TISSUE=Adrenal cortex;
RX MEDLINE-93342056; PubMed-8341687;
RA Lusson J., Vileau D., Hamelin J., Day R., Chretien M., Seidah N.G.;
RT "cDNA structure of the mouse and rat subtilisin/kexin-like PC5: a
RT candidate proprotein convertase expressed in endocrine and
RT nonendocrine cells."
RT Proc. Natl. Acad. Sci. U.S.A. 90:6691-6695(1993).
RN
[4]
RN PARTIAL SEQUENCE, AND SUBCELLULAR LOCATION.
RX MEDLINE-97103178; PubMed-8947550;
RA De Bie I., Marcinkiewicz M., Malide D., Lazure C., Nakayama K.,
RA Bendaayan M., Seidah N.G.;
RT "The isoforms of proprotein convertase PC5 are sorted to different
RT subcellular compartments."

```

```

RL J. Cell Biol. 135:1261-1275(1996).
[5]
RN DEVELOPMENTAL EXPRESSION.
RX MEDLINE-96293359; PubMed-8698813;
RA Constam D.B., Cailion M., Robertson E.J.;
RT "SPC4, SPC6, and the novel protease SPC7 are coexpressed with bone
RT morphogenetic proteins at distinct sites during embryogenesis."
RT J. Cell Biol. 134:181-191(1996).
RN
[6]
RN DEVELOPMENTAL EXPRESSION.
RX MEDLINE-97436919; PubMed-9291583;
RA Rancourt S.L., Rancourt D.E.;
RT "Murine subtilisin-like proteinase SPC6 is expressed during embryonic
RT implantation, somitogenesis, and skeletal formation."
RL Dev. Genet. 21:75-81(1997).
CC
CC -I- FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENDOPEPTASE ACTIVITY
CC OF CLEAVAGE AT THE RX(K/R) CONSENSUS MOTIF. MAY BE RESPONSIBLE
CC FOR THE MATURATION OF GASTROINTESTINAL PEPTIDES. MAY BE INVOLVED
CC IN THE CELLULAR PROLIFERATION OF ADRENAL CORTEX VIA THE ACTIVATION
CC OF GROWTH FACTORS.
CC
CC -I- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR
CC PROPROTEINS BY CLEAVAGE OF ARG-XAA-YAA-ARG-1-ZAA BONDS, WHERE XAA
CC CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
CC
CC -I- SUBCELLULAR LOCATION: PC5A IS SECRETED THROUGH THE REGULATED
CC SECRETORY PATHWAY. PC5B IS A TYPE I MEMBRANE PROTEIN LOCALIZED TO
CC A PARANUCLEAR POST-GOLGI NETWORK COMPARTMENT IN COMMUNICATION WITH
CC EARLY ENDOSOMES.
CC
CC -I- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; PC5B/LONG (SHOWN HERE)
CC AND PC5A/SHORT; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC
CC -I- TISSUE SPECIFICITY: PC5A IS EXPRESSED IN MOST TISSUES BUT IS MOST
CC ABUNDANT IN THE INTESTINE AND ADRENALS. PC5B IS EXPRESSED IN THE
CC INTESTINE, ADRENALS AND LUNG BUT NOT IN THE BRAIN.
CC
CC -I- DEVELOPMENTAL STAGE: WEAKLY EXPRESSED THROUGHOUT THE EMBRYO,
CC EXCEPT IN THE DEVELOPING NERVOUS SYSTEM, THE RIBS AND THE LIVER,
CC BUT MARKEDLY UPREGULATED AT DISCRETE SITES DURING DEVELOPMENT. AT
CC E6.5, PROMINENT EXPRESSION OBSERVED IN DIFFERENTIATED DECIDUA. AT
CC E7.5, INTENSE EXPRESSION IN EXTRAMEMBRYONIC ENDODERM, AMNION AND
CC NASCENT MESODERM. AT 8.5, ABUNDANT EXPRESSION IN SOMITES AND YOLK
CC SAC FOLLOWED BY A CONFINATION TO DERMATOTOME COMPARTMENT. BETWEEN
CC E9.5 AND E11.5, ABUNDANT EXPRESSION IN AER (THICKENED ECTODERMAL
CC CELLS OF LIMB BUDS). AT E12.5, EXPRESSION IN THE LIMBS IS CONFINED
CC TO THE CONDENSING MESENCHYM SURROUNDING THE CARILAGE. AT THIS
CC STAGE, STRONG EXPRESSION ALSO DETECTED IN VERTEBRAL AND FACIAL
CC CARTILAGE PRIMORDIA AND IN THE MUSCLE OF THE TONGUE. AT E16.5,
CC ABUNDANT EXPRESSION IN EPITHELIAL CELLS OF THE INTESTINAL VILLI.
CC ISOFORM A IS MOST ABUNDANT AT ALL STAGES BUT SIGNIFICANT LEVELS OF
CC ISOFORM B OCCUR AT E12.5.
CC
CC -I- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE
CC ASSISTING THE FOLDING OF THE ZYMOGEN WITHIN THE ENDOPLASMIC
CC RETICULUM.
CC
CC -I- DOMAIN: AC 1 AND AC 2 (CLUSTERS OF ACIDIC AMINO ACIDS) CONTAIN
CC SORTING INFORMATION. AC 1 DIRECTS TGN LOCALIZATION AND INTERACTS
CC WITH THE TGN SORTING PROTEIN PACS-1.
CC
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SUBTILASE FAMILY.
CC
CC -I- SIMILARITY: CONTAINS 1 HOMO B/P DOMAIN.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: D17583; BA004507.1; -
CC EMBL: D12619; BA002143.1; -
CC EMBL: L14932; AAA74636.1; -
CC PIR: JX0248; JX0248.
CC PIR: A48225; A48225.
CC HSSP: 099405; 1MPT.
CC MEROPS: S08.076; -.

```

```

DR MGD; MG1:97515; PCSK5.
DR InterPro: IPR000209; -
DR InterPro: IPR002884; -
DR Pfam: PF01483; P. 1.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
KW Hydrolase, Serine protease, Glycoprotein, Zymogen, Signal,
KW Cleavage on pair of basic residues, Repeat, Alternative splicing,
KW Transmembrane.
FT SIGNAL 1 34
FT PROPEP 35 116
FT CHAIN 117 1877
FT DOMAIN 117 1768
FT TRASMEM 1769 1789
FT DOMAIN 1790 1877
FT DOMAIN 117 452
FT DOMAIN 464 602
FT DOMAIN 638 1753
FT DOMAIN 1825 1844
FT DOMAIN 1856 1877
FT SITE 116 117
FT SITE 521 523
FT ACT_SITE 173 173
FT ACT_SITE 214 214
FT ACT_SITE 388 388
FT CARBOHYD 227 227
FT CARBOHYD 383 383
FT CARBOHYD 667 667
FT CARBOHYD 754 754
FT CARBOHYD 804 804
FT CARBOHYD 854 854
FT CARBOHYD 951 951
FT CARBOHYD 1016 1016
FT CARBOHYD 1220 1220
FT CARBOHYD 1317 1317
FT CARBOHYD 1523 1523
FT CARBOHYD 1711 1711
FT CARBOHYD 1733 1733
FT VARSPPLIC 878 915
FT VARSPPLIC 916 1877
FT VARSPPLIC 1877 AA; 209287 MM; EC850E2DF20EAI3 CRC64;
SQ SEQUENCE

```

```

AC P10041; Q9VDY2; Q99108;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROGENIC LOCUS DELTA PROTEIN PRECURSOR.
DL OR CG3619.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RC TISSUE-Embryo:
RT Vaessin H., Bremer K.A., Knust E., Campos-Ortega J.A.;
RT "The neurogenic gene Delta of Drosophila melanogaster is expressed in
RT neurogenic territories and encodes a putative transmembrane protein
RT with EGF-like repeats.";
RL EMBO J. 6:3431-3440(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-OREGON-R; TISSUE-Embryo:
RC MEDLINE=89196890; PubMed=3149249;
RT Kozczynski C.C., Alton A.K., Fehnel K., Kooh P.J., Muskavitch M.A.T.;
RT "Delta, a Drosophila neurogenic gene, is transcriptionally complex and
RT encodes a protein related to blood coagulation factors and epidermal
RT growth factor of vertebrates.";
RL Genes Dev. 2:1723-1735(1988).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY:
RC MEDLINE=20196006; PubMed=10731132;
RT Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RT Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RT George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RT Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
RT Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,
RT Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RT Abril J.F., Agbayani A., An H.-J., Andrews-Prankoch C., Baldwin D.,
RT Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RT Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RT Borokova D., Botchan M.R., Bouck J., Brokstein P., Botlier P.,
RT Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RT Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RT de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RT Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RT Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RT Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RT Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RT Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RT Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam G.,
RT Jalili M., Kalush F., Karpen G.H., Ke Z., Kennen J.A., Kethum K.A.,
RT Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RT Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RT Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
RT Mervilov G., Mishina N.V., Mobbarty C., Morris J., Moshrefi A.,
RT Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RT Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RT Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
RT Reiert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RT Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RT Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RT Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RT Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RT Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RT Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RT Gibbs R.A., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [4]
RP SEQUENCE OF 422-621 FROM N.A.
RC TISSUE-Embryo:

```

RX MEDLINE-87218537; PubMed-3107986;
 RA Knust E., Dietrich U., Tepass U., Bremer K.A., Weigelt D., Vaessin H.,
 RA Campos-Ortega J.A.;
 RT "Egf homologous sequences encoded in the genome of Drosophila
 RT melanogaster, and their relation to neurogenic genes.";
 RL EMO J. 6:761-766(1987).
 [5]
 RP PATTERN OF TRANSCRIPTION, AND CHARACTERIZATION.
 RX MEDLINE-91209246; PubMed-2128477;
 RA Haenlin M., Kratatschek B., Campos-Ortega J.A.;
 RT "The pattern of transcription of the neurogenic gene Delta of
 RT Drosophila melanogaster.";
 RL Development 110:905-914(1990).
 CC -1- FUNCTION: ESSENTIAL FOR PROPER DIFFERENTIATION OF ECTODERM. DL
 CC IS REQUIRED FOR THE CORRECT SEPARATION OF NEURAL AND EPIDERMAL
 CC CELL LINEAGES.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: DETECTED IN ALL AREAS WITH NEUROGENIC
 CC ABILITIES, FOR EXAMPLE THE NEUROGENIC ECTODERM AND THE PRIMORDIA
 CC OF THE SENSE ORGANS. LATER EXPRESSION IS RESTRICTED TO THOSE CELLS
 CC THAT HAVE ADOPTED A NEURAL FATE.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY.
 CC EXPRESSION IS HIGHEST EARLY IN EMBRYONIC DEVELOPMENT (STAGE 5) AND
 CC REDUCES TO A LOW LEVEL DURING LARVAL STAGES.
 CC -1- MISCELLANEOUS: SEPARATION OF NEUROBLASTS FROM THE ECTODERM INTO
 CC THE INNER PART OF EMBRYO IS ONE OF THE FIRST STEPS OF CNS
 CC DEVELOPMENT IN INSECTS, THIS PROCESS IS UNDER CONTROL OF THE
 CC NEUROGENIC GENES.
 CC -1- MISCELLANEOUS: NONCH AND SERRATE MAY INTERACT AT THE PROTEIN
 CC LEVEL, IT IS CONCEIVABLE THAT THE SERRATE AND DELTA PROTEINS MAY
 CC COMPLETE FOR BINDING WITH THE NOTCH PROTEIN.
 CC -1- SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE DELTA/SERRATE/JAGGED FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 CC EMBL: X06289; CAA29617.1; -
 CC EMBL: Y00222; CAA68369.1; -
 CC EMBL: AE003725; AAF55657.1; -
 CC EMBL: X05140; CAA28786.1; -
 CC PIR: S00670; S00670.
 CC PIR: A26637; A26637.
 CC HSSP: P00740; 11XA.
 CC FLYBASE: FBgn000463; DL.
 CC InterPro: IPR000152; -
 CC InterPro: IPR000561; -
 CC InterPro: IPR01774; -
 CC InterPro: IPR01881; -
 CC Pfam: PF01414; DSL; 1.
 CC Pfam: PF00008; EGF; 9.
 CC PROSITE: PS00010; ASX_HYDROXYL; 3.
 CC PROSITE: PS00022; EGF; 1; 9.
 CC PROSITE: PS01186; EGF; 2; 9.
 CC PROSITE: PS01187; EGF; CA; 2.
 CC Differentiation; Neurogenesis; Repeat; Transmembrane;
 CC EGF-like domain; Glycoprotein; Signal.
 CC EGF-LIKE 1
 FT CHAIN 1 18
 FT SIGNAL 1 18
 FT DOMAIN 19 833
 FT TRANSMEM 19 653
 FT TRANSMEM 654 833
 FT DOMAIN 678 833
 FT DOMAIN 227 258
 FT DOMAIN 256 289
 FT DOMAIN 291 329
 FT DOMAIN 331 372
 FT DOMAIN 374 416
 FT DOMAIN 418 451
 FT EGF-LIKE 6.

FT DOMAIN 453 489 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 491 527 EGF-LIKE 8.
 FT DISULFID 529 565 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
 FT DISULFID 231 240 BY SIMILARITY.
 FT DISULFID 235 246 BY SIMILARITY.
 FT DISULFID 248 257 BY SIMILARITY.
 FT DISULFID 260 271 BY SIMILARITY.
 FT DISULFID 266 277 BY SIMILARITY.
 FT DISULFID 279 288 BY SIMILARITY.
 FT DISULFID 295 307 BY SIMILARITY.
 FT DISULFID 301 317 BY SIMILARITY.
 FT DISULFID 319 328 BY SIMILARITY.
 FT DISULFID 335 348 BY SIMILARITY.
 FT DISULFID 342 360 BY SIMILARITY.
 FT DISULFID 362 371 BY SIMILARITY.
 FT DISULFID 378 388 BY SIMILARITY.
 FT DISULFID 383 404 BY SIMILARITY.
 FT DISULFID 406 415 BY SIMILARITY.
 FT DISULFID 422 433 BY SIMILARITY.
 FT DISULFID 427 439 BY SIMILARITY.
 FT DISULFID 441 450 BY SIMILARITY.
 FT DISULFID 457 468 BY SIMILARITY.
 FT DISULFID 462 477 BY SIMILARITY.
 FT DISULFID 479 488 BY SIMILARITY.
 FT DISULFID 495 506 BY SIMILARITY.
 FT DISULFID 500 515 BY SIMILARITY.
 FT DISULFID 517 526 BY SIMILARITY.
 FT DISULFID 533 544 BY SIMILARITY.
 FT DISULFID 538 553 BY SIMILARITY.
 FT DISULFID 555 564 BY SIMILARITY.
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 167 167 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 649 649 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 5 5 K -> N (IN REF. 1).
 FT CONFLICT 67 67 V -> L (IN REF. 1).
 FT CONFLICT 363 363 A -> R (IN REF. 1).
 FT CONFLICT 437 438 GK -> ET (IN REF. 3).
 FT CONFLICT 443 443 A -> S (IN REF. 1 AND 3).
 FT CONFLICT 459 459 G -> A (IN REF. 3).
 FT CONFLICT 490 490 S -> T (IN REF. 3).
 FT CONFLICT 591 591 T -> A (IN REF. 1 AND 3).
 FT CONFLICT 631 631 D -> N (IN REF. 1).
 FT CONFLICT 652 652 G -> A (IN REF. 1).
 FT CONFLICT 662 662 L -> M (IN REF. 1).
 FT CONFLICT 759 759 S -> T (IN REF. 1).
 FT CONFLICT 801 804 ACSS -> DLF1 (IN REF. 1).
 FT CONFLICT 812 812 V -> A (IN REF. 2).
 FT CONFLICT 815 817 SCA -> TD (IN REF. 1).
 FT CONFLICT 833 833 M -> YTPKIRKSMETIRNPHGADSTYYTKRLGWQNV
 RETPKMLLIEAV (IN REF. 1).
 SQ SEQUENCE 833 AA; 88840 MW; EDBA107A0003D9A1 CRC64;
 Query Match 10.4%; Score 93.5; DB 1; Length 833;
 Best Local Similarity 22.4%; Pred. No. 0.33;
 Matches 39; Conservative 12; Mismatches 66; Indels 57; Gaps 10;
 QY 9 PEEDRCSHS-----KLLPIMLMDSNKC---KVLQRENLAGEDSHLOEPALSCP 59
 DB 199 PRDSDGSHGTCSETGITLTLG--WQGDYCHLPKCA-----KGE-HGNDCKPQOCV 248
 QY 60 HMFEDRC-EGVCCKTPCKFDLIQHPKNSCFE-----CKESLETCCOKNKLFPNPTC- 111
 DB 249 QLGWKGALCNCEVLBNFCNCHGTGCKNPKYCTICENMGSLGICNDLVCTNHNPKCKNGTCE 308
 QY 112 -----SCEDR---CPNTRPCASG-----KTAACAKNC 135
 DB 309 NTGEGLYTCKCAPGYSGDDCENEIYSCDADVNPQNGGTCIDEPHTKTKYKCHC 362
 RESULT 11
 TENA_HUMAN


```

FT DISULFID 315 325 BY SIMILARITY.
FT DISULFID 319 330 BY SIMILARITY.
FT DISULFID 332 341 BY SIMILARITY.
FT DISULFID 346 356 BY SIMILARITY.
FT DISULFID 350 361 BY SIMILARITY.
FT DISULFID 363 372 BY SIMILARITY.
FT DISULFID 377 387 BY SIMILARITY.
FT DISULFID 381 392 BY SIMILARITY.
FT DISULFID 394 403 BY SIMILARITY.
FT DISULFID 408 418 BY SIMILARITY.
FT DISULFID 412 423 BY SIMILARITY.
FT DISULFID 425 434 BY SIMILARITY.
FT DISULFID 439 449 BY SIMILARITY.
FT DISULFID 443 454 BY SIMILARITY.
FT DISULFID 456 465 BY SIMILARITY.
FT DISULFID 470 480 BY SIMILARITY.
FT DISULFID 474 485 BY SIMILARITY.
FT DISULFID 487 496 BY SIMILARITY.
FT DISULFID 501 511 BY SIMILARITY.
FT DISULFID 505 516 BY SIMILARITY.
FT DISULFID 518 527 BY SIMILARITY.
FT DISULFID 532 542 BY SIMILARITY.
FT DISULFID 536 547 BY SIMILARITY.
FT DISULFID 549 558 BY SIMILARITY.
FT DISULFID 563 573 BY SIMILARITY.
FT DISULFID 567 578 BY SIMILARITY.
FT DISULFID 580 589 BY SIMILARITY.
FT DISULFID 594 604 BY SIMILARITY.
FT DISULFID 598 609 BY SIMILARITY.
FT DISULFID 611 620 BY SIMILARITY.
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 788 788 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1034 1034 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1079 1079 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1093 1093 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1119 1119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1184 1184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1210 1210 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1261 1261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1275 1275 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1301 1301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1366 1366 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1392 1392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1445 1445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1455 1455 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1485 1485 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1534 1534 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1809 1809 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2162 2162 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 1072 1435 MISSING (IN ISOFORM HT-5 AND ISOFORM

```

Query Match 10.4%; Score 93; DB 1; Length 2201;
 Best Local Similarity 26.1%; Pred. No. 0.82;
 Matches 41; Conservative 15; Mismatches 55; Indels 46; Gaps 11;

```

OY 10 EEDRC-----SHSKLCPIDIMDNKCKV---LQENPLAGTIDHSHLDEPALC 57
DB 357 EGGQVCDEGFGAGLDCSEKRPADC---HNRGRVGVRCEDDGTGA-DCGELKCPNQC 412
OY 58 GPHMMFDEDRC---ECVC-----KTPCKDLIOHPKNCSEFECKESLETCCKRKH 104
DB 413 SGH-----GRCVNGQVCDEGTGDEDCSLRCPNDC---HSRG-----RCVEGKVCVEGGRK 461
OY 105 LPHPTSCGDRCPHTRPCASGKTAC-----AKHCR 136
DB 462 GYDCSDMSPNDCHQHR-CVNGMKVCDDGTGTEGDR 497

```

RESULT 12

```

TRBM_MOUSE
ID TRBM_MOUSE STANDARD; PRT; 577 AA.
AC P15306;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE THROMBOMODULIN PRECURSOR (FETOMODULIN) (TM).
GN THBD.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=89128454; PubMed=2536925;
RA Dittman W.A., Majerus P.W.;
RT "Sequence of a cDNA for mouse thrombomodulin and comparison of the
RT predicted mouse and human amino acid sequences.";
RL Nucleic Acids Res. 17:802-802(1989).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=89008498; PubMed=2844823;
RA Dittman W.A., Kumada T., Sadler J.E., Majerus P.W.;
RT "The structure and function of mouse thrombomodulin. Phorbol
RT myristate acetate stimulates degradation and synthesis of
RT thrombomodulin without affecting mRNA levels in hemangioendothelial
RT J. Biol. Chem. 263:15815-15822(1988).
CC -1- FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR
CC THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS
CC COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE
CC ACTIVATED PROTEIN C (PROTEIN CA).
CC FACTOR VA AND FACTOR VIII, AND THEREBY REDUCES THE AMOUNT OF
CC THROMBIN GENERATED.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING
CC THROMBOMODULIN (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sb.ch/announce/
CC or send an email to license@isb-sb.ch).
CC -----
DR EMBL: X14432; CAA32597.1; -
DR PIR: A32001; A32001.
DR PIR: S08488; S08488.
DR HSP: P07204; 1EGT.
DR MGD: MGI:98736; Thbd.
DR InterPro: IPR000152; -
DR InterPro: IPR000561; -
DR InterPro: IPR001304; -
DR InterPro: IPR001491; -
DR InterPro: IPR001881; -
DR Pfam: PF00008; EGF 4.
DR Pfam: PF00059; Lectin_c.1.
DR PRINTS: PR00907; THROMBOMODULIN.
DR PROSITE: PS00010; ASX_HYDROXYL; 2.
DR PROSITE: PS00022; EGF_1; FALSE_NEG.
DR PROSITE: PS01186; EGF_2; 3.
DR PROSITE: PS01187; EGF_CA; 2.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
KW Endothelial cell; Receptor; Blood coagulation; Repeat; Transmembrane;
KW Glycoprotein; Signal; -
FT CHAIN 1 577 THROMBOMODULIN.
FT DOMAIN ? 517 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 518 541 POTENTIAL.
FT DOMAIN 542 577 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 240 280 EGF-LIKE 1.

```



```

FT DOMAIN 283 323 EGF-LIKE 2.
FT DOMAIN 324 362 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 364 404 EGF-LIKE 4.
FT DOMAIN 403 438 EGF-LIKE 5.
FT DOMAIN 440 480 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT DISULFID 244 255 BY SIMILARITY.
FT DISULFID 251 264 BY SIMILARITY.
FT DISULFID 266 279 BY SIMILARITY.
FT DISULFID 287 295 BY SIMILARITY.
FT DISULFID 291 307 BY SIMILARITY.
FT DISULFID 309 322 BY SIMILARITY.
FT DISULFID 328 339 BY SIMILARITY.
FT DISULFID 335 348 BY SIMILARITY.
FT DISULFID 350 361 BY SIMILARITY.
FT DISULFID 368 377 BY SIMILARITY.
FT DISULFID 373 387 BY SIMILARITY.
FT DISULFID 389 403 BY SIMILARITY.
FT DISULFID 407 416 BY SIMILARITY.
FT DISULFID 412 424 BY SIMILARITY.
FT DISULFID 426 438 BY SIMILARITY.
FT DISULFID 444 454 BY SIMILARITY.
FT DISULFID 449 463 BY SIMILARITY.
FT DISULFID 465 479 BY SIMILARITY.
FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 408 408 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 494 494 O-LINKED (GLYCOSAMINOGLYCAN) (BY
SEQUENCE 577 AA; 61867 MW; B20E50B0FE745014 CRC64;
SIMILARITY).

```

Query Match 10.3%; Score 92.5; DB 1; Length 577;
 Best Local Similarity 24.7%; Pred. No. 0.3; Mismatches 43; Conservative 15; Indels 67; Gaps 11;

```

OY 11 EDRCSHSKRLCP-IDMLDMSKC-KCVLOEENPLAGTEDSHLOEP---ALC--GPHM 61
DB 261 EPRC-----LCPRDMDLQADRGSCARPVVQSCNELCEHFCVSNABVGSVSCMCEGYOL 315
OY 62 MEPEDECEVC--CK---TPCKPLDIQHPKNCSCF-----ECKESLETCCQKHKL 106
DB 316 AAGHRCEDVDKCGKGNPCPOLCVNTKGGFECFCYDGYELVDGCVLLDPCFGSN--- 372
OY 107 HPDTSCEDRCFPHTRP-----CASG-----KTACAKHC 135
DB 373 -----CFCQCPVSTDYRCICAFKPFKPKCEMFCNETSCPADC 416

RESULT 13
ITBX_DROME
ID ITBX_DROME STANDARD: PRT; 846 AA.
AC P11584;
DT 01-OCT-1989 (rel. 12, Created)
DT 01-OCT-1989 (rel. 12, Last sequence update)
DT 15-JUL-1998 (rel. 36, Last annotation update)
DE INTEGRIN BETA-SUBUNIT MYOSPHEROID PRECURSOR.
GN L(1)MYS.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Anthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88190122; PubMed=3128792;
RA Mackrell A.J., Blumberg B., Haynes S.R., Fessler J.H.;
RT "The lethal myospheroid gene of Drosophila encodes a membrane protein
homologous to vertebrate integrin beta subunits."
RL Proc. Natl. Acad. Sci. U.S.A. 85:2633-2637(1988).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=94163982; PubMed=8119134;

```

```

RA Grinblat Y., Zusman S., Yee G., Hynes R.O., Kafatos F.C.;
RT "Functions of the cytoplasmic domain of the beta PS integrin subunit
during Drosophila development."
RL Development 120:91-102(1994).
CC CC -1- FUNCTION: NOT KNOWN. PROBABLY PLAYS A ROLE IN CELL ADHESION.
CC CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR EMBL; J03251; AAA28714.1; -.
DR PIR; A30889; A30889.
DR FLYBASE; FBgn0004657; mys.
DR InterPro; IPR000561; -.
DR InterPro; IPR001169; -.
DR InterPro; IPR002369; -.
DR Pfam; PF00362; Integrin_B; 1.
DR PROSITE; PS00243; INTEGRIN_BETA; 3.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
KW Integrin; Cell adhesion; Transmembrane; Glycoprotein; Repeat;
KW Extracellular matrix; Cytoskeleton; Signal.
FT SIGNAL 1 23
FT CHAIN 24 846
FT DOMAIN 24 776
FT TRANSMEM 777 799
FT DOMAIN 800 846
FT DOMAIN 115 143
FT DOMAIN 307 687
FT REPEAT 507 560
FT REPEAT 561 605
FT REPEAT 606 646
FT REPEAT 647 687
FT CARBOHYD 72 72
FT CARBOHYD 266 266
FT CARBOHYD 277 277
FT CARBOHYD 403 403
FT CARBOHYD 428 428
FT CARBOHYD 557 557
FT CARBOHYD 603 603
FT CARBOHYD 644 644
FT CARBOHYD 718 718
SQ SEQUENCE 846 AA; 92687 MW; 56249715F0F99A1D CRC64;

```

Query Match 10.3%; Score 92.5; DB 1; Length 846;
 Best Local Similarity 27.1%; Pred. No. 0.41; Mismatches 39; Conservative 13; Indels 55; Gaps 11;

```

OY 10 EPRCS-HSKRLCPIDMLDMSKCVCVLOEENPLAGTEDSHLOEPALCGPHMFADE 65
DB 518 QANSCSGHGTSMGICNCDSYFNKCEC-----SADLTSKRANDTSCADSI 569
OY 66 D--RCEVC-----KTPCKPLDIQHPKNCSC--FECKSELTCCOKHKLFPD--TCS 112
DB 570 DCSGRHCVCAGACECHKRPDIETIS-GKHCECDNFSCERNRQLCS-----GPDHGTCE 623
OY 113 CEDRCFPHTRPFCASGKTACAKHCR 136
DB 624 C-GRC-----KCRPGWTGSGNCGQ 641

RESULT 14
FUR2_DROME
ID FUR2_DROME STANDARD: PRT; 1680 AA.
AC P30432;
DT 01-APR-1993 (rel. 25, Created)

```

DT 01-APR-1993 (Rel. 25, last sequence update)
 DT 15-JUL-1998 (Rel. 36, last annotation update)
 DE FURIN-LIKE PROTEASE 2 PRECURSOR (EC 3.4.21.75) (FURIN 2).
 GN FUR2.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92381036; PubMed=1512259;
 RA Roebroek A.J.M., Creemers J.W.M., Paul I.G.L., Kurzik-Dumke U.,
 RA Rantrop M., Gateff E.A.F., Leunissen J.A.M., van de Ven W.J.;
 RT "Cloning and functional expression of Dfurn2, a subtilisin-like
 RT proprotein processing enzyme of Drosophila melanogaster with multiple
 RT repeats of a cysteine motif";
 RL J. Biol. Chem. 267:17208-17215(1992).
 CC -1- FUNCTION: FURIN IS LIKELY TO REPRESENT THE UBIQUITOUS ENDOPEPTIDASE
 CC ACTIVITY WITHIN CONSTITUTIVE SECRETORY PATHWAYS AND CAPABLE OF
 CC CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR
 CC PROPEPTIDES BY CLEAVAGE OF ARG-XAA-YAA-ARG-1-ZAA BONDS, WHERE XAA
 CC CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS. RELEASES ALBUMIN,
 CC COMPLEMENT COMPONENT C3 AND VON WILLEBRAND FACTOR FROM THEIR
 CC RESPECTIVE PRECURSORS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
 CC SUBTILASE FAMILY. HIGH SIMILARITY WITH OTHER FURIN-LIKE ENZYMES.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; M94375; AAA28551.1; -
 CC PIR; A43434; A43434.
 CC DR HSSP: O99405; 1MP.
 CC DR Elybase; FBgn0004598; Fur2.
 CC DR InterPro: IPR000209; -
 CC DR InterPro: IPR002174; -
 CC DR InterPro: IPR002884; -
 CC DR Pfam: PF00757; Furin-like; 1.
 CC DR Pfam: PF01483; p; 1.
 CC DR Pfam: PF00082; Peptidase_S8; 1.
 CC DR PRINTS; PR00723; SUBTILISIN.
 CC DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 CC DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 CC DR PROSITE; PS00138; SUBTILASE_SER; 1.
 CC KM Hydrolyase; Serine protease; Glycoprotein; Signal; Transmembrane;
 CC MultiGene family; Zymogen.
 CC FT SIGNAL 1 ?
 CC FT PROPEP 319
 CC FT CHAIN 320 1680
 CC FT ACT_SITE 418 418
 CC FT ACT_SITE 457 457
 CC FT ACT_SITE 638 638
 CC FT DOMAIN 962 1444
 CC FT REPEAT 962 1007
 CC FT REPEAT 1008 1057
 CC FT REPEAT 1058 1104
 CC FT REPEAT 1105 1153
 CC FT REPEAT 1154 1205
 CC FT REPEAT 1206 1254
 CC FT REPEAT 1255 1299
 CC FT REPEAT 1300 1346
 CC FT REPEAT 1347 1393
 CC FT REPEAT 1394 1444
 CC FT TRANSMEM 1508 1532
 CC FT DOMAIN 1533 1680
 CC CARBOHYD 3
 CC
 CC POTENTIAL.
 CC CYTOPLASMIC (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC

FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 928 928 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1061 1061 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1182 1182 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1275 1275 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1278 1278 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1440 1440 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1680 AA; 183599 MW; 0A99CE8770A8E293 CRC64;
 QY 11 EDRSHSKKLPIDMLDSNKC-----KVLQENPLACT-----EDSHLQEPALCP 59
 DB 956 ETRGDGKKILHSCDAECDSGCGYRGPTQVACSHYRLDWTQVSRCPFRFPNOVGICWP 105
 QY 60 HMFEDRC-----CVKTPCPKDLIOHPKNCSECKESL 96
 DB 1016 ----CHDTCEFCAGAGPDSCLTCAPAHLHYIDLAVLCLOFCPDGFEFNSRRNRCVPCPNC 1071
 QY 97 ETC-----CQKHLFHPDNCSCEDRCPTH-----RPSAGKACAKHCFPREK 141
 DB 1072 ASCDHPDYCTCDHLVMEH--KCYSCAPLDYTERDNKCAFCSTCAT-CNGPDD 1127
 RESULT 15
 ID SELP_BOVIN STANDARD; PRT; 402 AA.
 AC P49907; O19003;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, last sequence update)
 DT 01-OCT-2000 (Rel. 40, last annotation update)
 DE SELENOPROTEIN P-LIKE PROTEIN PRECURSOR.
 GN SELP OR SELP.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_Taxid=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cerebellum;
 RX MEDLINE=95364621; PubMed=7637580;
 RA Saijoh K., Saito N., Lee M.J., Fujii M., Kobayashi T., Sumino K.;
 RT "Molecular cloning of cDNA encoding a bovine selenoprotein P-like
 RT protein containing 12 selenocysteines and a (His-Pro) rich domain
 RT insertion, and its regional expression";
 RL Brain Res. Mol. Brain Res. 30:301-311(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98019090; PubMed=9358058;
 RA Fujii M., Saijoh K., Kobayashi T., Fujii S., Lee M.J., Sumino K.;
 RT "Analysis of bovine selenoprotein P-like protein gene and availability
 RT of metal responsive element (MRE) located in its promoter.";
 RL Gene 199:211-217(1997).
 CC -1- FUNCTION: IT CONSTITUTES A MAJOR SELENIUM POOL IN THE BRAIN AND
 CC MAY PLAY AN IMPORTANT ROLE IN DEVELOPING AND/OR MODULATING THE
 CC MORPHOLOGY OF NEURONS AND/OR GLIAL CELLS.
 CC -1- TISSUE SPECIFICITY: BRAIN AND KIDNEY. MOST PROMINENTLY EXPRESSED
 CC IN THE CEREBELLAR CORTEX, HIPPOCAMPUS AND OLFACTORY BULB.
 CC -1- MISCELLANEOUS: THE SELENOCYSTEINES ARE ALL ENCODED BY THE OPAL
 CC CODON, UGA.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D25220; BAA04949.1; -;
DR EMBL; D88033; BAA23414.1; -;
DR EMBL; D88031; BAA23414.1; JOINED.
DR EMBL; D88032; BAA23414.1; JOINED.
DR HSSP; P04355; 4MT2.
KW Glycoprotein; Signal; Selenium; Selenocysteine; Repeat.
FT SIGNAL 1 19
FT CHAIN 20 402
FT SE_CYS 59 59
FT SE_CYS 297 297
FT SE_CYS 307 307
FT SE_CYS 338 338
FT SE_CYS 350 350
FT SE_CYS 363 363
FT SE_CYS 365 365
FT SE_CYS 372 372
FT SE_CYS 388 388
FT SE_CYS 390 390
FT SE_CYS 397 397
FT SE_CYS 399 399
FT DOMAIN 204 239
FT DOMAIN 260 266
FT CONFLICT 178 181
FT CONFLICT 256 256
FT CONFLICT 282 282
FT CONFLICT 312 312
SQ SEQUENCE 402 AA; 45018 MW; B7CF18751FB08EFF CRC64;
H-P REPEATS.
POLY-HIS.
SRPQ -> KALE (IN REF. 2).
T -> P (IN REF. 2).
L -> V (IN REF. 2).
Y -> D (IN REF. 2).

Query Match 10.38; Score 92; DB 1; Length 402;
Best Local Similarity 24.58; Pred. No. 0.25;
Matches 38; Conservative 11; Mismatches 46; Indels 60; Gaps 9;

QY 38 EENPLAGTSDHSHLQEPALCGPHMFEDEDECEC-----VCKTPCKDLI- 81
DB 252 ENPPTSGLHHHHRHK-----GPOROGSHDNCDPPLGSESLQPSLPQKLCRKRCINOLLC 307
QY 82 QHPK-----NCSC-----FECKESLETCCOKHKLHPDTCSCD-----RC 117
DB 308 QPPTYSSEALSSCCCHRLVFEKTSALTCCQCEKL--PSLCSGGLAEENVIESCQC 365
QY 118 PFHTRPC-----ASGKTAC--AKHCRFP 138
DB 366 RLPPAACGAAGQOLNPTEASTKSCSKNKAHKCKP 400

Search completed: October 17, 2001, 14:51:51
Job time: 477 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:50:58 ; Search time 128.06 seconds
(without alignments)
158.072 Million cell updates/sec

Title: US-09-427-657-4_COPY_202_354
Perfect score: 897
Sequence: 1 IIRRSIQIPEEDRCSHSHKTL.....HCRFPKERRAAGPMSRRNP 153

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_16:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp-organelle:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp_unclassified:*
13: sp-vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	897	100.0	354	4 043915	043915 homo sapien
2	739.5	82.4	358	11 P97946	P97946 mus musculu
3	595	66.3	326	11 035251	035251 rattus norv.
4	466	52.0	122	6 0961X1	0961X1 bos taurus
5	233.5	26.0	420	6 09X550	09X550 bos taurus
6	231	25.8	418	13 057352	057352 coturnix co
7	166	18.5	1698	5 094438	094438 chironomus
8	150	16.7	1704	5 094446	094446 chironomus
9	116.5	13.0	1187	2 049549	049549 mycoplasma
10	112.5	12.5	648	5 09NKD7	09NKD7 drosophila
11	112.5	12.3	701	5 09VJ04	09VJ04 drosophila
12	110.5	12.3	704	3 09VZF2	09VZF2 drosophila
13	108.5	12.1	757	5 09VZF2	09VZF2 drosophila
14	108	12.0	748	6 010741	010741 bos taurus
15	107.5	12.0	765	5 09NL50	09NL50 sarcophaga
16	106	11.8	830	4 043701	043701 homo sapien
17	106	11.8	830	4 04162	04162 homo sapien
18	105.5	11.8	814	11 090Y12	090Y12 mus musculu
19	105	11.7	2515	5 024551	024551 drosophila

20	105	11.7	2515	5 09VNU6	09VNU6 drosophila
21	105	11.7	2731	5 061307	061307 drosophila
22	105	11.7	2731	5 018366	018366 drosophila
23	103.5	11.5	815	11 088839	088839 mus musculu
24	103	11.5	2946	5 018857	018857 caenorhabd
25	102.5	11.4	800	5 09VB78	09VB78 drosophila
26	102.5	11.4	837	5 09NAS7	09NAS7 anopheles g
27	102.5	11.4	3680	5 09VRO8	09VRO8 drosophila
28	101.5	11.3	747	10 09LGM8	09LGM8 oryza sativ
29	100.5	11.2	1114	11 09JRW7	09JRW7 mus musculu
30	100.5	11.2	2100	5 021281	021281 caenorhabd
31	99	11.0	1322	5 09NAT0	09NAT0 anopheles g
32	98.5	11.0	5376	11 088799	088799 mus musculu
33	97.5	10.9	776	6 09XSL6	09XSL6 macaca fasc
34	97.5	10.9	925	4 09H318	09H318 homo sapien
35	97.5	10.9	1036	4 09NZV1	09NZV1 homo sapien
36	97.5	10.9	1599	5 009983	009983 caenorhabd
37	97.5	10.9	3843	5 09U5D0	09U5D0 drosophila
38	97.5	10.9	3843	5 09VU94	09VU94 drosophila
39	97	10.8	386	6 09N2H6	09N2H6 bos taurus
40	96.5	10.8	336	1 027597	027597 methanobact
41	96.5	10.8	349	1 059575	059575 methanobact
42	96.5	10.8	788	13 007012	007012 xenopus lae
43	96.5	10.8	2910	11 055225	055225 mus musculu
44	96.5	10.8	2920	11 09ROM0	09ROM0 mus musculu
45	96	10.7	377	5 09V392	09V392 drosophila

ALIGNMENTS

RESULT 1
ID 043915 PRELIMINARY; PRT; 354 AA.
AC 043915;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE GROWTH FACTOR FIGF.
GN FIGF OR VEGF-D.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98140120; PubMed=9479493;
RA Rocchianni M., Lestingi M., Luddi A., Orlandini M., Franco B.,
RA Rossi E., Ballabio A., Zuffardi O., Oliviero S.;
RT "Human FIGF: cloning, gene structure, and mapping to chromosome Xp22.1
RT between the FIGA and the GPR genes.";
RL Genomics 47:207-216(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RX MEDLINE=97349118; PubMed=9205122;
RA Yamada Y., Nezu J., Shimane M., Hirata Y.;
RT "Molecular cloning of a novel vascular endothelial growth factor,
RT VEGF-D.";
RL Genomics 42:483-488(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98118549; PubMed=9435229;
RA Achen M.G., Jeltsch M., Kuk E., Maekinen T., Vitall A., Wilks A.F.,
RA Allitalo K., Stacker S.A.;
RT "Vascular endothelial growth factor D (VEGF-D) is a ligand for the
RT tyrosine kinases VEGF receptor 2 (Flk1) and VEGF receptor 3 (Flt4).";
RL Proc. Natl. Acad. Sci. U.S.A. 95:548-553(1998).
DR EMBL: Y12864; CAA73371.1; JOINED.
DR EMBL: Y12865; CAA73371.1; JOINED.
DR EMBL: Y12866; CAA73371.1; JOINED.
DR EMBL: Y12867; CAA73371.1; JOINED.
DR EMBL: Y12868; CAA73371.1; JOINED.

DR EMBL: Y12869; CAA73371.1; JOINED.
DR EMBL: Y12870; CAA73371.1; JOINED.
DR EMBL: D89630; BAA24264.1; -.
DR EMBL: AJ000185; CAA03942.1; -.
DR EMBL: Y12863; CAA73370.1; -.
DR HSSP: P15692; 1VPP.
DR InterPro: IPR000072; -.
DR Pfam: PF00341; PDGF; 1.
DR ProDom: PD001629; -; 1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS50278; PDGF_2; 1.
DR SMART: SM00141; PDGF; 1.
SQ SEQUENCE 354 AA; 40444 MW; 2048D769D735173E CRC64;

Query Match 100.0%; Score 897; DB 4; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.2e-86;
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IIRRSIQIPEDRCSHSKKLCPIIDMLDSNKKCVLQENRPLAGTEDHSHLOEPALCGPH 60
DB 202 IIRRSIQIPEDRCSHSKKLCPIIDMLDSNKKCVLQENRPLAGTEDHSHLOEPALCGPH 261
QY 61 MMEDEDRCEVCCTPCPKDIIQHPKNSCFECKESLETCCQKHKLHNPDTSCSEDRCRPFH 120
DB 262 MMEDEDRCEVCCTPCPKDIIQHPKNSCFECKESLETCCQKHKLHNPDTSCSEDRCRPFH 321
QY 121 TRPCASGKTACAKHCRPFKERRAAGPHSRKPN 153
DB 322 TRPCASGKTACAKHCRPFKERRAAGPHSRKPN 354

RESULT 2
P97946 PRELIMINARY; PRT; 358 AA.
ID P97946;
AC P97946;
DT 01-MAY-1997 (TEMBLrel. 03, Created)
DT 01-MAY-1997 (TEMBLrel. 03, Last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR D (C-FOS INDUCED GROWTH FACTOR).
GN VEGF-D OR FIGF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=97030254; PubMed=8876195;
RA Ordanini M., Marconini L., Ferruzzi R., Oliviero S.;
RT "Identification of a c-fos-induced gene that is related to the
RT platelet-derived growth factor/Vascular endothelial growth factor
RT family";
RL Proc. Natl. Acad. Sci. U.S.A. 93:11675-11675(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RX MEDLINE=97349118; PubMed=9205122;
RA Yamada Y., Nezu J., Shimane M., Hirata Y.;
RT "Molecular cloning of a novel vascular endothelial growth factor,
RT VEGF-D";
RT Genomics 42:483-488(1997).
DR EMBL: X99572; CAA67892.1; -.
DR EMBL: D89628; BAA14002.1; -.
DR HSSP: P15692; 1VPP.
DR MGI:108037; Figf.
DR InterPro: IPR000072; -.
DR Pfam: PF00341; PDGF; 1.
DR ProDom: PD001629; -; 1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS50278; PDGF_2; 1.
DR SMART: SM00141; PDGF; 1.
SQ SEQUENCE 358 AA; 40908 MW; 6636B17FBF07037C CRC64;

Query Match 82.4%; Score 739.5; DB 11; Length 358;
Best Local Similarity 82.4%; Pred. No. 4.1e-70;
Matches 126; Conservative 9; Mismatches 17; Indels 1; Gaps 1;

QY 1 IIRRSIQIPEDRCSHSKKLCPIIDMLDSNKKCVLQENRPLAGTEDHSHLOEPALCGPH 60
DB 207 IIRRSIQIPEDRCSHSKKLCPIIDMLDSNKKCVLQENRPLAGTEDHSHLOEPALCGPH 266
QY 61 MMEDEDRCEVCCTPCPKDIIQHPKNSCFECKESLETCCQKHKLHNPDTSCSEDRCRPFH 120
DB 267 MMEDEDRCEVCCTPCPKDIIQHPKNSCFECKESLETCCQKHKLHNPDTSCSEDRCRPFH 326
QY 121 TRPCASGKTACAKHCRPFKERRAAGPHSRKPN 153
DB 327 TRPCASGKTACAKHCRPFKERRAAGPHSRKPN 358

RESULT 3
O35251 PRELIMINARY; PRT; 326 AA.
ID O35251;
AC O35251;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR D.
GN VEGF-D.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE DAWLEY;
RX MEDLINE=97349118; PubMed=9205122;
RA Yamada Y., Nezu J., Shimane M., Hirata Y.;
RT "Molecular cloning of a novel vascular endothelial growth factor,
RT VEGF-D";
RT Genomics 42:483-488(1997).
DR EMBL: AF014827; AAB6557.1; -.
DR HSSP: P15692; 1VPP.
DR InterPro: IPR000072; -.
DR Pfam: PF00341; PDGF; 1.
DR ProDom: PD001629; -; 1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS50278; PDGF_2; 1.
DR SMART: SM00141; PDGF; 1.
SQ SEQUENCE 326 AA; 37112 MW; 1261AFA373596C00 CRC64;

Query Match 66.3%; Score 595; DB 11; Length 326;
Best Local Similarity 88.3%; Pred. No. 5.7e-55;
Matches 98; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 IIRRSIQIPEDRCSHSKKLCPIIDMLDSNKKCVLQENRPLAGTEDHSHLOEPALCGPH 60
DB 207 IIRRSIQIPEDRCSHSKKLCPIIDMLDSNKKCVLQENRPLAGTEDHSHLOEPALCGPH 266
QY 61 MMEDEDRCEVCCTPCPKDIIQHPKNSCFECKESLETCCQKHKLHNPDTSC 111
DB 267 MMEDEDRCEVCCTPCPKDIIQHPKNSCFECKESLETCCQKHKLHNPDTSC 317

RESULT 4
O9GLX1 PRELIMINARY; PRT; 122 AA.
ID O9GLX1;
AC O9GLX1;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR-D (FRAGMENT).
GN VEGF-D.

OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Mandiata S.J., Pepper M.S.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF099135; AAG29747.1; -
FT NON_TER 1 1
FT 122 122
SQ SEQUENCE 122 AA; 13820 MW; CC504B00E29D54EB CRC64;

Query Match 52.0%; Score 466; DB 6; Length 122;
Best Local Similarity 90.9%; Pred. No. 9.3e-42;
Matches 80; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 IIRRSIQIPEDRCSSKRLCPIDMLDMSNKCVCVLOEENPLAGTEHSHLOEPALCGPH 60
DB 35 IIRRSIQIPEDRCSSKRLCPVDMWDSNKCVCVLOEENPLAGMEDHTLOELALCGH 94
QY 61 MFEDRCCECVCKTPCKDLIQHNCS 88
DB 95 MKFEDRCCECVCKTPCKDLIQHNCS 122

RESULT 5
Q9XS50 PRELIMINARY; PRT; 420 AA.
AC Q9XS50;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEART;
RA Liu X., Yonekura H., Yamagishi S., Yamamoto Y., Yamamoto H.;
RT "Structure and expression of bovine VEGF family.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB004275; BAA77687.1; -
DR HSSP; P15692; IVP.
DR InterPro: IPR000072; -
DR Pfam: PF00341; PDGF_1;
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS50278; PDGF_2; 1.
DR SMART: SM00141; PDGF; 1.
KW Signal.
FT SIGNAL 1 20
FT CHAIN 21 420
SQ SEQUENCE 420 AA; 46681 MW; 58BA84317A3CEB2D CRC64;

Query Match 26.0%; Score 233.5; DB 6; Length 420;
Best Local Similarity 28.5%; Pred. No. 6.5e-17;
Matches 53; Conservative 25; Mismatches 43; Indels 65; Gaps 8;

QY 1 IIRRSIQIPEDRCSSKRLCPIDMLDMSNKCVCVLOEENPLAGTED----- 47
DB 225 IIRRSIQIPEDRCSSKRLCPIDMLDMSNKCVCVLOEENPLAGTED----- 47
QY 48 -----HSHLOE-----PALCGPHMFEDRCCE 69
DB 282 GPNKELDETCOCVCKGLOASSCGPHKELDRSCOCVCKKNKLFPSGCGANREDFENTCO 341

QY 70 CVCKTPCKDLIQHNCSCEFEKESLETCCOKHKLFPDPSCEDRCFPHTRPCASGKT 129
DB 342 CVCKTPCKDLIQHNCSCEFEKESLETCCOKHKLFPDPSCEDRCFPHTRPCASGKT 129
QY 130 ACARHC 135
DB 391 NRVKHC 396

RESULT 6
O57352 PRELIMINARY; PRT; 418 AA.
AC O57352;
ID O57352;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR.
GN VEGF-C.
OS Colurnix colurnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Colurnix.
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98167900; PubMed=9435294;
RA Eichmann A., Corbel C., Jaffredo T., Breant V., Joukov V., Kumar V.,
RA Altalo K., Le Douarin N.M.;
RT "Avian VEGF-C: cloning, embryonic expression pattern and stimulation
RT of the differentiation of VEGFR2-expressing endothelial cell.
RT precursors.";
RT RT
RL Development 125:743-752(1998).
DR EMBL; Y15837; CAA75799.1; -
DR HSSP; P15692; IVP.
DR InterPro: IPR000072; -
DR InterPro: IPR002400; -
DR Pfam: PF00341; PDGF; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR PRODOM; PD001629; -; 1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS50278; PDGF_2; 1.
DR SMART: SM00141; PDGF; 1.
KW Signal.
FT SIGNAL 1 31
FT CHAIN 111 418
SQ SEQUENCE 418 AA; 46839 MW; 099BFC79151BF2B CRC64;

Query Match 25.8%; Score 231; DB 13; Length 418;
Best Local Similarity 29.0%; Pred. No. 1.2e-16;
Matches 54; Conservative 21; Mismatches 59; Indels 52; Gaps 6;

QY 1 IIRRSIQIPEDRCSSKRLCPIDMLDMSNKCVCVLOEENPLAGTED----- 47
DB 223 IIRRSIQIPEDRCSSKRLCPIDMLDMSNKCVCVLOEENPLAGTED----- 47
QY 48 -----HSHLOE-----PALCGPHMFEDRCCEV 71
DB 282 NKELEDETCOCVCKGVRPISCGPHKELDRASCCOMCKNKLFPSCGPNKEFDEKOCV 341

QY 72 CKTPCKDLIQHNCSCEFEKESLETCCOKHKLFPDPSCEDRCFPHTRPCASGKT 130
DB 342 CKTPCKDLIQHNCSCEFEKESLETCCOKHKLFPDPSCEDRCFPHTRPCASGKT 130

QY 131 CANHCR 136
DB 401 ABEVCR 406

RESULT 7
O94438 PRELIMINARY; PRT; 1698 AA.
ID O94438;
AC O94438;

DT 01-FEB-1997 (Tremblrel. 02, Created)
DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE 185 KPA SILK PROTEIN.
GN SP185.
OS Chironomus pallidivittatus (Midge).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomidae; Chironomidae; Chironominae; Chironomus.
OX NCBI_TaxID=7151;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPERMATOPHYTES;
RA Case S.T., Cox C., Bell W.C., Hoffman R.T., Martin J., Hamilton R.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U54640; AAA99803.1; -
DR InterPro: IPR000561; -
DR InterPro: IPR000853; -
DR PRINTS; PR00876; MTNEMATODE.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
SQ SEQUENCE 1698 AA; 186164 MW; 82AF0838D9E2360E CRC64;

Query Match 18.5%; Score 166; DB 5; Length 1698;
Best Local Similarity 26.9%; Pred. No. 2,4e-09;
Matches 50; Conservative 21; Mismatches 51; Indels 64; Gaps 11;
QY 10 EEDRSHSKLCPIDM-----LMDSNKCCVLOENPLAGTEDHSHLOE----- 53
DB 239 DDSCACK---CPVNMKKKECIAXKINPNPNCCEGCAELNPOKKPKKECECECK 255
QY 54 -PALGPHMFEDRCRC-----VCKTP-----CPKDLIOHP----- 84
DB 296 ILAKCNDGOVFNKPCSCNPGASADVKCKPQIYNNESSCGCPVDM-QKPIDGCPRPQ 354
QY 85 -----KNCSCCECKSLSTCCGKHKLFPDPCSCEDRCFPHTRPASCRTACAKHC----- 135
DB 355 RWESDSCSC-EC-PLKQDCKNGKVMATYICQ--ICPRDALPCAAKRGESCGCKCI 409
QY 136 -REFPE 140
DB 410 NREPK 415

RESULT 8
ID Q94446 PRELIMINARY; PRT: 1704 AA.
AC Q94446;
DT 01-FEB-1997 (Tremblrel. 02, Created)
DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE 220 KDA SILK PROTEIN.
GN SP220.
OS Chironomus thummi thummi (Midge).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomidae; Chironomidae; Chironominae; Chironomus.
OX NCBI_TaxID=7155;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SALIVARY GLAND;
RA Case S.T., Cox C., Bell W.C., Hoffman R.T., Martin J., Hamilton R.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U54641; AAA99804.1; -
DR InterPro: IPR000561; -
DR InterPro: IPR000853; -
DR PRINTS; PR00876; MTNEMATODE.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
SQ SEQUENCE 1704 AA; 185745 MW; 3A3F20247C8F1E28 CRC64;

Query Match

16.7%; Score 150; DB 5; Length 1704;

Best Local Similarity 24.2%; Pred. No. 1.2e-07;
Matches 46; Conservative 27; Mismatches 37; Indels 80; Gaps 11;
QY 14 CSHSKLCPIDM-----CPIDMLDMSKCCVLOENPLAGTEDHSHLOEPA- 55
DB 1216 CSASOKYSNVTSCGCGNPGAPAKCGPNOYWCNCCVCPK-----NMKPAD 1264
QY 56 LCGPHMFEDRCRCVCTPCP-----KDLIOHPKN-CSCECKESLET--CCOKHKLFP 108
DB 1265 NCG-NKMNNDKACCECKPCGPPEAGCKGVOKMKNNTAC-ECPRGKATPASCDDKSMNP 1322
QY 109 DTGSCEDR-----CPH-----TRPCASGT----- 129
DB 1323 DSCSCCKSKMPGGCPSNOOWNCETCKCECSGTOTCPAGOSMDSOTCCSPATGCKTG 1382
QY 130 ---ACAKHCR 136
DB 1383 AQFWCAKCK 1392

RESULT 9
ID Q49549 PRELIMINARY; PRT: 1187 AA.
AC Q49549;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
DE REPEAT REGIONS IN POTENTIAL METAL BINDING PROTEIN GENE REGION.
OS Mycoplasma hyorhins.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscites;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2100;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GD1.
RX MEDLINE=95014025; PubMed=7928953;
RA Deng G., McIntosh M.A.;
RT "An amplifiable DNA region from the Mycoplasma hyorhins genome";
RL J. Bacteriol. 176:5929-5937(1994).
DR EMBL; L11447; AAA6228.1; -
SQ SEQUENCE 1187 AA; 135027 MW; 66A82AE50EDE93E CRC64;

Query Match 13.0%; Score 116.5; DB 2; Length 1187;
Best Local Similarity 26.8%; Pred. No. 0.00029;
Matches 42; Conservative 20; Mismatches 62; Indels 33; Gaps 8;
QY 14 CSHSKLCPIDMLDMSKCCVLOENPLAGTEDH---SHLOE-PALGPHMFEDRC 69
DB 293 CDICKNLEHSEILYRLKNGVVMLELETETELHYADSPVQKPEKCGSLKETESCD 352
QY 70 C-VCKTPCPKDLIOHPKNCSEFE-----CKSLSTCCGKHKLFPDPCSC----- 113
DB 353 CEACK--CQ-----ECENSCSELTCGCEATCSCAOEHCCEOCESCAPNTTACTEEN 406
QY 114 -----EDRCFPHTRPASCAGTAC---AKHCFPKPKRA 143
DB 407 CECTESTCGCENPCCEBEACDCSEHCCEVDETA 443

RESULT 10
ID Q9NKD7 PRELIMINARY; PRT: 648 AA.
AC Q9NKD7;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE HYPOTHETICAL 69.4 KDA PROTEIN.
GN BG:DS00180.14 OR CG18146.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.


```

OX NCBL_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=, CN BW SP;
RA MEDLINE=99403001; PubMed=10471707;
RA Ashburner M., Mista S., Rhee J., Lewis S.E., Blazet R., Davis T.,
RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
RA Palazzolo M., Reese M.G., Spreading A., Tsang G., Wan K., Whitehead K.,
RA Celinker S., Rubin G.M.;
RT "An exploration of the sequence of a 2.9-Mb region of the genome of
RT Drosophila melanogaster: the Ach region.";
RL Genet. 153:179-219(1999).
[12]
RN SEQUENCE FROM N.A.
RP STRAIN=, CN BW SP;
RC Celinker S.E., Agdayani A., Arcalata T.T., Baxter E., Blazet R.G.,
RA Butenhof C., Chame M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA Houston K.A., Hummasti S.R., Karia K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lomotan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeita A.,
RA Sethi H., Shit E., Syrtkas R.R., Wan K.H., Weinburg T., Zhang R.,
RA Zieran L.L., Rubin G.M.;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF003408; AAF44845.1;
DR Flybase: FBgn0028939; BG:DS00180.14.
DR InterPro: IPR000561;
DR PROSITE: PS00186; EGF_2; UNKNOWN_6.
DR SMART: SM00181; EGF; 1.
KM Hypothetical protein.
SQ SEQUENCE 648 AA; 69419 MW; 3F16E0BFA94CF8 CRC64;

Query Match
Best Local Similarity 12.5%; Score 112.5; DB 5; Length 648;
Matches 51; Conservative 16; Mismatches 68; Indels 97; Gaps 13;

OY 9 PEEDRC--SHSKL-----CPIDMLMDSNKKCKVL---QENPLA-GIEDHS 49
DB 143 PNCRCRHPGHEORLGVWICDPICSSGCANGYCGAEVACCKGVANKNLTLASGE--- 199
OY 50 HLOEPALGCP-----HMFDEDERCCEV--CKTPCPDLIOHPKNCSC 89
DB 200 -----PVCNRPCTNGTCISPGHCACSEGHVFAEGSHNECVPSRSCSENGYCSPPRCCEC 254
OY 90 FE-----CKESLET--CCQKHKLFHPDTCSCF-----DRC--PFHTRPCASG- 127
DB 255 HEEFEKTSFPHRCSPTRPCGCGNSRCAPDTCADVGYYFVNGSTTECEFCPRCNRNGI 314
OY 128 -----KTCAKHCPRPKKRAAGPHSRKNP 153
DB 315 CSSPGVCTCLEGFQALLSFYCIPIVCSKTCIHGSCVAPNRCRFTG--YRPNP 364

RESULT 11
OYVU4 PRELIMINARY; PRT; 701 AA.
AC O9VU4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE CG18146 PROTEIN.
GN BG:DS00180.14 OR CG18146.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta:
OC Pterygota: Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha:
OC Ephydroidea: Drosophilidae: Drosophila.
OX NCBL_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;

```

```

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazet R.G., Chame M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
RA Abrell J.F., Agdayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodet A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Moadary C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA Switz R., Tector C., Turner R., Venter G.E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AF003642; AAF53366.1;
DR HSSP: F02876; 9WGA.
DR Flybase: FBgn0028939; BG:DS00180.14.
DR InterPro: IPR000561;
DR PROSITE: PS00186; EGF_2; UNKNOWN_6.
DR SMART: SM00181; EGF; 1.
SQ SEQUENCE 701 AA; 75366 MW; 95159C04C9C09BD4 CRC64;

Query Match
Best Local Similarity 12.5%; Score 112.5; DB 5; Length 701;
Matches 51; Conservative 16; Mismatches 68; Indels 97; Gaps 13;

OY 9 PEEDRC--SHSKL-----CPIDMLMDSNKKCKVL---QENPLA-GIEDHS 49
DB 196 PNCRCRHPGHEORLGVWICDPICSSGCANGYCGAEVACCKGVANKNLTLASGE--- 252
OY 50 HLOEPALGCP-----HMFDEDERCCEV--CKTPCPDLIOHPKNCSC 89
DB 253 -----PVCNRPCTNGTCISPGHCACSEGHVFAEGSHNECVPSRSCSENGYCSPPRCCEC 307
OY 90 FE-----CKESLET--CCQKHKLFHPDTCSCF-----DRC--PFHTRPCASG- 127
DB 308 HEEFEKTSFPHRCSPTRPCGCGNSRCAPDTCADVGYYFVNGSTTECEFCPRCNRNGI 367
OY 128 -----KTCAKHCPRPKKRAAGPHSRKNP 153
DB 368 CSSPGVCTCLEGFQALLSFYCIPIVCSKTCIHGSCVAPNRCRFTG--YRPNP 417

RESULT 12
OYVU4 PRELIMINARY; PRT; 704 AA.
AC O74567;
DT 01-NOV-1998 (TREMBLrel. 08, Created)

```

DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DE 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DT Q174 PROTEIN.
 GN Q174.
 OS Trichoderma harzianum.
 OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Trichoderma.
 NC NCBL_TaxID=5544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CECT 2413;
 RX MEDLINE=98263335; PubMed=9600944;
 RA Rey M., Ohno S.A., Plator-Toro J.A., Jose A., Ilobell A., Benitez T.;
 RT "Unexpected homology between inducible cell wall protein Q174 of
 filamentous fungi and BR3 salivary protein of the insect Chironomus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:6212-6216(1998).
 DR EMBL: X95671; CA64974.1; -
 DR InterPro: IPR000561; -
 DR PROSITE: PS01186; EGF_2; 1.
 SQ SEQUENCE 704 AA; 77925 MW; 63414BDEEC365EBC CRC64;

Query Match 12.3%; Score 110.5; DB 3; Length 704;
 Best Local Similarity 23.8%; Pred. No. 0.00081;
 Matches 39; Conservative 13; Mismatches 51; Indels 61; Gaps 9;
 QY 14 CSHSKKLCPTIDMLMSNKKCVL-QEENPLAGTETHSHLOEPALCGPHMDEDECEVC 72
 DB 389 CS-----CPDQGYWNGKOCACPYGQIFNGKQCVPD-----CGMDATDYNNQKKVC 434
 QY 73 KT-----PCPKDLIQHPKNCSC-----FECKESLENC-----CG 101
 DB 435 KNGEYFDOKTKTCSCPDGQYWGKQCVCPYKIFDGKQCVGDADHEDNNQKKVCN 494
 QY 102 KHKLFHPD--TSCED-----RCPFHRCASGKTACAKKC 135
 DB 495 KNEIITDSKSKTSCPDQYWGKQCVCPYKIFDGKQCVGDADHEDNNQKKVCN 533

RESULT 13
 Q9VZF2 PRELIMINARY; PRT; 757 AA.
 AC Q9VZF2;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE CG15011 PROTEIN.
 GN CG15011.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Ephydroida; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidae; Drosophilidae; Drosophila.
 NC NCBL_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Gelinkler S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Abdayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borzova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durlin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Mostrel A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhao M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003480; AAF47871.1; -
 DR FlyBase: FBgn0035518; CG15011.
 DR InterPro: IPR000561; -
 DR InterPro: IPR000967; -
 DR InterPro: IPR001368; -
 DR Pfam: PF01422; zf-NF-X1; 7.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
 DR SMART: SM00438; zNF_NFX; 1.
 SQ SEQUENCE 757 AA; 84556 MW; 34A7F3C28BD6428F CRC64;

Query Match 12.1%; Score 108.5; DB 5; Length 757;
 Best Local Similarity 25.7%; Pred. No. 0.0014;
 Matches 39; Conservative 14; Mismatches 60; Indels 39; Gaps 9;
 QY 13 RCS--HKKKLCPTIDMLMSNKKCVLQENPLAGTETHSHLOE-----PALCGPHMPD 64
 DB 253 KQNVCHQKCPKPSKSLQPCCE--QRESKVMVNSDKRMKQNVNCGAFPGAGLHI--- 307
 QY 65 EDRCEVCYCT-----PCPKDLIQHP--KNCSCFECKESLENC-----CQKRLFHPDTC 111
 DB 308 ---CEKVCHAGRCQDCEPLQVRSQPCGKNCNCGDQPCPEKICGKQLSCNKR----- 357
 QY 112 SCEDRCPFHRCASGKTACAKHCRPKKRA 143
 DB 358 -COSVC--HNGPCYCKLESQINCRCGKTKS 386

RESULT 14
 Q10741 PRELIMINARY; PRT; 748 AA.
 AC Q10741;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE DISINTEGRIN-METALLOPROTEASE PRECURSOR (PC 3.4.24.-) (MYELIN-ASSOCIATED
 DE METALLOPROTEINASE) (MADP).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 NC NCBL_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=BRAIN;
 RX MEDLINE=96276398; PubMed=8694785;
 RA Howard L., Lu X., Mitchell S., Griffiths S., Glynn P.;
 RT "Molecular cloning of MADP, a catalytically active mammalian
 disintegrin-metalloprotease expressed in various cell types.";
 RL Biochem. J. 317:45-50(1996).
 CC -i- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

DR EMLL: AB031297; BAA02143.1; -.
DR InterPro: IPR000561; -.
DR PROSITE: PS01186; EGF_2; UNKNOWN_7.
DR SMART: SM00181; EGF_1.
KW Signal.
FT SIGNAL 1 18 POTENTIAL.
SQ SEQUENCE 765 AA; 83909 MW; FF25BEF1288FE007 CRC64;

Query Match	12.0%;	Score 107.5;	DB 5;	Length 765;
Best Local Similarity	22.6%;	Pred. No. 0.0018;		
Matches	33;	Conservative	8;	Mismatches 60;
				Indels 45;
				Gaps 6;

```

0Y 11 EDRSHSKKLCPIIDMLDMSNCKCVLQENPLAGTEDESHLOEPALCGPMHMFEDRCEC 70
Db 295 EPT---KNGCENGFCSEPGKCEC-----HEGYETEPHLSP----- 329
0Y 71 VCKTPCEKDLQHHKNSCFE-----CKESLETCCOKHKLPHPTSCED---- 115
Db 330 ICKEPCVNGICAPDPTCKMNGYIADTKEVCEPYCPKCKGNKNCIAPGVCTONPQYS 389
0Y 116 ----RC-PFHTRPCASGKTACAKHCR 136
Db 390 LFFHCIPICIKPCLOGTCTAPDTCR 415

```

Search completed: October 17, 2001, 14:50:59
Job time: 425 sec

Query Match	12.0%	Score 108;	DB 6;	Length 748;
Best Local Similarity	22.3%	Pred. No. 0.0016;		
Matches 37;	Conservative 18;	Mismatches 53;	Indels 58;	Gaps 8;

```

QY 12 DRCRSHK-KLCP---IMLWMSNCKCVLQLEENLTACTEDSHSLQEPALCGPHMFEEDR 67
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 425 DKLNKKRSLCSIKINISQVLEKKRNNCFVESGP-----ICGNHVEQGE 470
QY 68 CEC-----VCKTPCEKDLIOHKNSC-----FECKSELTCCQKHKLPHDPTCSD-- 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 471 CDCGSDCKDECCYDANQ-PEGKCKKLPRKQCSPPQGFECSTAHCAFKSKTEKCRDSD 529
QY 116 -----RCPEFHPPCASGRTA-----CAKH 134
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 530 CAKEBICMGITALCPASDPKRNPTDCNRHRYQVCLNGCCAGSICEKH 575

```

RESULT	15			
09NLT50				
ID	09NLT50	PRELIMINARY;	PRT;	765 AA.
AC	09NLT50:			
DT	01-OCT-2000 (TIREMBLrel. 15, Created)			
DT	01-OCT-2000 (TIREMBLrel. 15, Last sequence update)			
DT	01-MAR-2001 (TIREMBLrel. 16, Last annotation update)			
DE	120-KDA PROTEIN PRECURSOR.			
OS	Sarcophaga peregrina (Flesh fly) (Boettcherisca peregrina).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Oestridae; Sarcophagidae; Sarcophaga.			
OX	NCBI_TaxID=7386;			
RN	{1}			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=HEMOCYTE;			
RA	Kobayashi A., Hori S., Natoli S.;			
RT	"Sarcophaga prepupal hemocyte-specific 120-kDa protein.";			
RL	Submitted (AUG-1999) to the EMBL/GenBank/DBU databases.			

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:46:02 ; Search time 115.93 seconds
(without alignments)
57.000 Million cell updates/sec

Title: US-09-427-657-4_COPY_93_201
Perfect score: 598
Sequence: 1 FVDETLKVIDEWMQRTGCS.....KVANHCTCKCLPRAPRHYS 109

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues
Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq_0601.*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	598	100.0	109	20	AAV23889
2	598	100.0	109	21	AA11931
3	598	100.0	325	19	AAW53240
4	598	100.0	325	22	AAV97572
5	598	100.0	354	19	AAW49036
6	598	100.0	354	19	AAW53241
7	598	100.0	354	19	AAW4293
8	598	100.0	354	21	AA10649
9	598	100.0	354	21	AA129049
10	598	100.0	354	21	AAV70750
11	598	100.0	354	21	AAV70983

12	598	100.0	354	22	AAV97573	Human VEGF-D1 prot
13	598	100.0	354	22	AA137606	Human VEGF-D, Hom
14	582	97.3	354	22	AA170685	Human VEGF-D, Hom
15	582	97.3	620	18	AAW14994	Human c-Fos induce
16	579	96.8	326	19	AAW44296	Rat VEGF-D
17	575	96.2	321	19	AAW53243	Mus musculus endot
18	575	96.2	337	20	AAV08286	Mus musculus endot
19	575	96.2	358	18	AAW14992	Human growth facto
20	575	96.2	358	19	AAW53242	Murine c-Fos induc
21	575	96.2	358	19	AAW44295	Mus musculus vascu
22	532	89.0	178	20	AAV08287	Mouse VEGF-D
23	377	63.0	113	20	AAV08285	Human growth facto
24	377	63.0	307	20	AAW86222	Human VEGF-C trunc
25	377	63.0	318	20	AAV08284	Human growth facto
26	377	63.0	350	16	AAW82686	Vascular endotheli
27	377	63.0	350	20	AAV30519	Vascular endotheli
28	377	63.0	350	20	AAV22321	A truncated Vascu
29	377	63.0	350	21	AAV97145	Truncated human VE
30	377	63.0	350	22	AAV97577	Human VEGF-2 prote
31	377	63.0	399	20	AAW86237	Human VEGF-C full
32	377	63.0	419	18	AAW17837	Human foetal liver
33	377	63.0	419	18	AAW00932	Human Flt4 recepto
34	377	63.0	419	18	AAW11478	Human VEGF-D
35	377	63.0	419	19	AAW75740	Human VEGF-D
36	377	63.0	419	20	AAV30518	Human VEGF-D
37	377	63.0	419	20	AAV22320	Full length human
38	377	63.0	419	20	AAW86203	Human VEGF-D
39	377	63.0	419	21	AAW10648	Human VEGF protein
40	377	63.0	419	21	AAW29048	Human VEGF-C prote
41	377	63.0	419	21	AAV97144	Vascular endotheli
42	377	63.0	419	21	AAV70749	Human prepro-vascu
43	377	63.0	419	21	AAV70982	Human VEGF-B prote
44	377	63.0	419	22	AAV97570	Human VEGF-B prote
45	377	63.0	419	22	AAW37605	Human VEGF-C, Hom

ALIGNMENTS

RESULT 1	
AAV23889	standard; Protein: 109 AA.
ID	AAV23889;
AC	AAV23889;
DT	21-SEP-1999 (first entry)
DE	Human vascular endothelial growth factor (VEGF)-D.
XX	Vascular endothelial growth factor; VEGF; VEGF-D; malignant melanoma;
XX	tumour; psoriasis; angiogenesis; lymphangiogenesis; skin graft;
XX	wound healing; lymphedema; scleroderma; anhydrotic ectodermal dysplasia.
OS	Homo sapiens.
XX	MO9933485-A1.
PN	08-JUL-1999.
XX	23-DEC-1998; 98WO-US27373.
XX	29-MAY-1998; 98US-0087392.
PR	24-DEC-1997; 97AU-0001131.
XX	(LUDW-) LUDWIG INST CANCER RES.
PA	Achen MG, Altitalo K, Stacker SA;
XX	WPI, 1999-405368/34.
DR	A human cell line stably expressing vascular endothelial growth
XX	factor D, useful for treating melanomas or tumours expressing VEGF-D
PT	
XX	

PS Claim 6; Page 72; 79pp; English.
XX The present sequence represents human vascular endothelial growth
CC factor (VEGF)-D. The specification describes a human cell line
CC which stably expresses VEGF-D, or fragments/analogs having VEGF-D
CC biological activity. VEGF-D antagonists, e.g. antisense nucleic acids
CC or triplex DNA, VEGF-D variants or antibodies (especially chimeric
CC antibodies), are useful for the treatment or alleviation of malignant
CC melanomas, tumours or psoriasis. Angiogenesis and lymphangiogenesis
CC stimulating amounts of VEGF-D can be administered to enhance the
CC acceptance and/or healing of skin grafts or to stimulate the healing
CC of a surgical or traumatic wound to the skin. Lymphangiogenesis
CC stimulating amounts of VEGF-D can be used to treat lymphedema.
CC Endothelial proliferation stimulating amounts of VEGF-D are used to
CC treat scleroderma. Vascularisation stimulating amounts of VEGF-D can
CC be used to treat anhydrotic ectodermal dysplasia. VEGF-D antibodies are
CC useful for detecting tumours expressing VEGF-D. Fully-processed VEGF-D
CC can be used to stimulate at least one VEGF-D bioactivity chosen from
CC endothelial cell proliferation, migration, survival and differentiation
CC and lymphangiogenesis without inducing vascular permeability.
XX
XX Sequence 109 AA:
SQ
Query Match 100.0%; Score 598; DB 20; Length 109;
Best Local Similarity 100.0%; Pred. No. 3.5e-58;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FYDIEFLKVIDEMORTGCSPRETCVEVASELGSNTFFKPPCVNVFRCGCCNEESLI 60
Db 1 fydielelkyvideawqrtqcspretcvevaseelgkstntffkppcvnvfrtgcgcneesi 60
OY 61 CMMTSTSYISKOLFETISVPLTSPVLPVVKVANHGTCKCLPTAPRRHPS 109
Db 61 cmmstcsyslksqfletsypeltsvpelvpvkvanhtgckclptaprrhps 109
RESULT 2
AAB11931
ID AAB11931 standard; protein; 109 AA.
XX
XX AAB11931;
XX
XX 20-NOV-2000 (first entry)
XX
XX Human truncated VEGF-D.
XX
XX Truncated VEGF-D: vascular endothelial growth factor; human;
KW monoclonal antibody; VEGF receptor; VEGFR-2; VEGFR-3;
KW vascular permeability disorder; endothelial cell proliferative disorder;
KW angiogenic disorder; lymphangiogenic disorder;
KW neovascularisation disorder; endothelial cell differentiation disorder;
KW cancer; diabetic retinopathy; psoriasis; arthropathy; pulmonary oedema;
KW detection; diagnosis; imaging; lymphatic vasculature.
XX
XX Homo sapiens.
XX
XX WO200037025-A2.
XX
XX 29-JUN-2000.
XX
XX 21-DEC-1999; 99WO-US31332.
XX
XX 21-DEC-1998; 98US-0113254.
XX 17-MAY-1999; 99US-0134556.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Achen MG, Stacker SA.
XX
XX WPI; 2000-442498/38.
XX
XX Novel compositions comprising antibodies reactive to vascular

PT endothelial growth factor-D, useful for treating, e.g. angiogenesis,
PT Lymphangiogenesis and neovascularization disorders -
XX
XX Claim 1; Fig 1; 44pp; English.
XX
XX This sequence represents a 109 amino acid truncated human VEGF-D
CC (vascular endothelial growth factor D), lacking both the N- and
CC C-terminal regions. The invention relates to a monoclonal antibody, or
CC fragments thereof, which is specifically reactive with the truncated
CC human VEGF-D, and methods of preparing the antibody. The antibody of the
CC invention interferes with the binding of VEGF-D to the VEGF receptors
CC VEGFR-2 and VEGFR-3, but does not interfere with the binding of VEGF to
CC these receptors and additionally is not reactive with VEGF-C. The
CC antibody may be used to treat disorders associated with vascular
CC permeability, endothelial cell proliferation, angiogenesis,
CC lymphangiogenesis, neovascularisation and endothelial cell
CC differentiation, especially cancer, diabetic retinopathy, psoriasis, and
CC arthropathies. The antibody may also be used to treat fluid accumulation
CC in the heart and/or lung via modulation of vascular permeability. It may
CC additionally be used to detect VEGF-D and may be used to image lymphatic
CC vasculature in tissue.
XX
XX Sequence 109 AA:
SQ
Query Match 100.0%; Score 598; DB 21; Length 109;
Best Local Similarity 100.0%; Pred. No. 3.5e-58;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FYDIEFLKVIDEMORTGCSPRETCVEVASELGSNTFFKPPCVNVFRCGCCNEESLI 60
Db 1 fydielelkyvideawqrtqcspretcvevaseelgkstntffkppcvnvfrtgcgcneesi 60
OY 61 CMMTSTSYISKOLFETISVPLTSPVLPVVKVANHGTCKCLPTAPRRHPS 109
Db 61 cmmstcsyslksqfletsypeltsvpelvpvkvanhtgckclptaprrhps 109
RESULT 3
AAM53240
ID AAM53240 standard; protein; 325 AA.
XX
XX AAM53240;
XX
XX 03-AUG-1998 (first entry)
XX
XX Homo sapiens vascular endothelial growth factor D (VEGF-D).
XX
XX vascular endothelial growth factor; VEGF-D; angiogenesis;
KW modification; acceleration; wound healing; tissue; organ;
KW transplants; collateral circulation; infarction; arterial stenosis;
KW coronary artery disease; inhibition; cancer; treatment;
KW diabetic retinopathy; lung disorders; blood circulation;
KW gaseous exchange; chronic obstructive airway disease;
KW intestinal malabsorptive syndrome; biopsy; metastatic risk;
KW detection; diagnosis; congestive heart failure.
XX
XX Homo sapiens.
XX
XX
XX
XX Key Location/Qualifiers
FH 126..128
FT /note="potential N-linked glycosylation site"
FT 156..158
FT /note="potential N-linked glycosylation site"
FT 258..260
FT /note="potential N-linked glycosylation site"
XX
XX WO9807832-A1.
XX
XX 26-FEB-1998.
XX
XX 21-AUG-1997; 97WO-US14696.
XX

PR 01-JUL-1997; 97US-0051426.
 PR 23-AUG-1996; 96AU-0001825.
 PR 23-AUG-1996; 96US-0023751.
 PR 11-NOV-1996; 96AU-0003554.
 PR 14-NOV-1996; 96US-0031097.
 PR 05-FEB-1997; 97AU-0004954.
 PR 10-FEB-1997; 97US-0038814.
 PR 19-JUN-1997; 97AU-0007435.
 XX (LUDM-) LUDMIG INST CANCER RES.
 PA (UYHE-) UNIV HELSINKI LICENSING LTD.
 XX Achen MG, Altalo K, Stackel SA, Wilks AF;
 PI WPI; 1998-179057/16.
 DR N-PSDB; AAV20806.
 XX New isolated vascular endothelial growth factor-D - used to develop
 PT products for use in e.g. modifying angiogenesis or treating lung,
 PT heart or intestinal disorders
 XX Claim 16; Pages 57-58; 101pp; English.
 XX The sequence is that of human breast vascular endothelial growth factor
 CC D (VEGF-D). VEGF-D can be used for e.g. acceleration of angiogenesis
 CC in wound healing, tissue or organ transplantation, or to establish
 CC collateral circulation in tissue infarction or arterial stenosis,
 CC such as coronary artery disease, and inhibition of angiogenesis in
 CC the treatment of cancer or of diabetic retinopathy. It can also be
 CC used in the treatment of lung disorders to improve blood circulation
 CC in the lung and/or gaseous exchange between the lungs and the blood
 CC stream or to improve blood circulation to the heart and O2 gas
 CC permeability in cases of cardiac insufficiency, to improve blood
 CC flow and gaseous exchange in chronic obstructive airway disease,
 CC or to treat malabsorptive syndromes in the intestinal tract.
 CC Quantitation of VEGF-D in cancer biopsy specimens may be useful
 CC as an indicator of future metastatic risk. Antagonists can be used
 CC for treating e.g. conditions such as congestive heart failure,
 CC involving accumulations of fluid in the lung resulting from
 CC increases in vascular permeability. The products can also be used
 CC for detection and diagnosis.
 CC XX
 S0 Sequence 325 AA;
 Query Match 100.0%; Score 598; DB 19; Length 325;
 Best Local Similarity 100.0%; Pred. No. 1.3e-57;
 Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FYDIETLKVIDEEMORTQCSPRETCVEVASSELGKSTNTFFKPPCVAVFRGGCCNEESLI 60
 DB 64 FYDIETLKVIDEEMWRTQCSPRETCVEVASSELGKSTNTFFKPPCVAVFRGGCCNEESLI 123
 QY 61 CMMTSTSYISKOLFETISVPLTSVPELVKPVKVAHNTGCKCLPTPRRPHYS 109
 DB 124 CMMTSTSYISKOLFETISVPLTSVPELVKPVKVAHNTGCKCLPTPRRPHYS 172
 RESULT 4
 ID AAY97572
 XX AAY97572 standard; Protein; 325 AA.
 AC AAY97572;
 XX 05-APR-2001 (first entry)
 DT Human VEGF-D protein sequence.
 DE Human VEGF-D protein sequence.
 XX Human: angiogenic protein; wound healing; vascular tissue repair;
 KM peripheral arterial disease; critical limb ischaemia; coronary disease;
 KM angiogenesis; tumour; inflammation; diabetic retinopathy; psoriasis;
 KM rheumatoid arthritis; autoimmune disease; allergy; cancer; therapy;
 KM infectious disease; neurodegeneration;

KM Vascular endothelial growth factor-D; VEGF-D.
 XX Homo sapiens.
 OS WO200075163-A1.
 PN 14-DEC-2000.
 PD 01-JUN-2000; 2000MO-US14925.
 PF 03-JUN-1999; 99US-0137796.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Ruben SM, Hu J, Cao L;
 PI WPI; 2001-071057/08.
 DR N-PSDB; AAV91006.
 XX New nucleic acid encoding angiogenic proteins, useful e.g. for
 PT promoting healing of wounds and treating peripheral arterial disease,
 PT critical limb ischaemia or coronary disease -
 XX Claim 11; Page 226-227; 244pp; English.
 XX This sequence is vascular endothelial growth factor-D (VEGF-D),
 CC which is an angiogenic protein of the invention. The angiogenic proteins
 CC and the DNA sequences encoding them, are used to prevent, treat or
 CC ameliorate disease and to detect diseases, or susceptibility, by
 CC detecting mutations or the presence or amount of angiogenic protein
 CC expression. Particularly they are used to stimulate wound healing,
 CC growth of damaged bone and tissue, and for repair of vascular tissue,
 CC especially peripheral arterial disease, critical limb ischaemia or
 CC coronary disease. Antagonists of the sequences are used to inhibit
 CC angiogenesis in tumours and to treat inflammation (where associated with
 CC increased vascular permeability), diabetic retinopathy, rheumatoid
 CC arthritis or psoriasis. Agonists are also useful for stimulating
 CC (lymph)angiogenesis. The proteins are also used to identify specific
 CC binding agents (potential therapeutic agents) and to raise antibodies.
 CC The antibodies are useful as therapeutic (ant)agonists; for detection,
 CC purification and targeting of proteins for in vivo or in vitro diagnosis
 CC (including imaging) or for therapy (including when linked to e.g. a label
 CC or cytotoxin); and for immunotyping of cells; e.g. for detecting minimal
 CC residual disease or haematopoietic progenitor/stem cells. It is also
 CC contemplated that the sequences might be useful for treating a very wide
 CC range of other disorders, e.g. autoimmune diseases; allergy; cancer;
 CC infectious diseases (viral, bacterial, fungal or parasitic);
 CC neurodegeneration, also as chemotactic agents or for stimulating
 CC regeneration of the nervous system etc.
 CC XX
 S0 Sequence 325 AA;
 Query Match 100.0%; Score 598; DB 22; Length 325;
 Best Local Similarity 100.0%; Pred. No. 1.3e-57;
 Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FYDIETLKVIDEEMORTQCSPRETCVEVASSELGKSTNTFFKPPCVAVFRGGCCNEESLI 60
 DB 64 FYDIETLKVIDEEMWRTQCSPRETCVEVASSELGKSTNTFFKPPCVAVFRGGCCNEESLI 123
 QY 61 CMMTSTSYISKOLFETISVPLTSVPELVKPVKVAHNTGCKCLPTPRRPHYS 109
 DB 124 CMMTSTSYISKOLFETISVPLTSVPELVKPVKVAHNTGCKCLPTPRRPHYS 172
 RESULT 5
 ID AAM49036
 XX AAM49036 standard; Protein; 354 AA.
 AC AAM49036;
 XX 26-OCT-1998 (first entry)
 DT

```

XX DE Human zveg12 growth factor.
XX KM Human zveg12 growth factor; mitogen; fibroblast; smooth muscle cell;
XX KM venous stasis ulcer; diabetic ulcer; skin wound; chemotactic effect;
XX KM angiogenic effect; tumour; diabetic retinopathy; psoriasis; arthritis;
XX KM scleroderma.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Peptide 1..23 /note="Signal peptide"
XX FT Peptide 24..108 /note="Pro-region"
XX FT Binding-site /note="Receptor binding domain"
XX FT Region 206..256 /note="Cysteine-rich domain"
XX FT Region 257..274 /note="Halbani ring motif"
XX FT Region 275..294 /note="Halbani ring motif"
XX FT Region 295..354 /note="Cysteine-rich domain"
XX FT Region /note="Cysteine-rich domain"
XX PN MO9824811-A2.
XX PD 11-JUN-1998.
XX PE 20-NOV-1997; 97MO-US20888.
XX PR 18-SEP-1997; 97US-0933455.
XX PR 06-DEC-1996; 96US-0759657.
XX PA (ZYMO) ZYMOGENETICS INC.
XX PI Conklin DC, Gilbert T, Hart CE, Nygaard S, Sheppard PO;
XX DR WPI: 1998-333256/29.
XX DR N-PSDB; AAV32823.
XX PT New isolated vascular endothelial growth factor - used to develop
XX PT products for treating e.g. wounds, burns, myocardial infarction,
XX PT tumours, psoriasis, arthritis, restenosis or organ transplants
XX PS Claim 1; Pages 53-54; 77pp; English.
XX CC The present sequence represents a human zveg12 growth factor encoded
XX CC by the zveg12 cDNA which was isolated from a human heart cDNA library.
XX CC zveg12 protein in a dimeric form acts as a mitogen for fibroblasts or
XX CC smooth muscle cells. zveg12 is claimed to be useful for stimulating the
XX CC revascularisation of tissue or the re-endothelialisation of vascular
XX CC tissue. zveg12 is particularly claimed to be useful for the treatment
XX CC of full-thickness skin wounds, including venous stasis ulcers and
XX CC diabetic ulcers. The zveg12 protein is also claimed to be useful as an
XX CC additive in tissue adhesives for promoting revascularisation of the
XX CC healing tissue. Antagonists against zveg12 can be used to block its
XX CC mitogenic, chemotactic and angiogenic effects. The antagonists may
XX CC therefore be useful for reducing growth of solid tumours by inhibiting
XX CC neovascularisation of the developing tumour or by directly blocking
XX CC tumour cell growth, in the treatment of diabetic retinopathy, psoriasis,
XX CC arthritis, and scleroderma.
XX SO Sequence 354 AA:

```

```

Query Match 100.0%; Score 598; DB 19; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.4e-57;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 FYDIEFLKVIDEWMORTQSPRETCEVASELCKSTNTEFFKPCVAVFRCGGCNEESLI 60
|||||

```

```

Db 93 fydielkvideewgrtqspretcevaselgkstinlffkpcvavfrcggcneesli 152
Qy 61 CMNTSTYSISKOLFETISVPLTSVPELVKVAHNGCKCLPTAPRRPYS 109
Db 153 cmntstysiskqlfetsvpltsvpelevkvanhtgckclptaprrpys 201
RESULT 6
ID AAW53241
AAW53241 standard; Protein; 354 AA.
AC AAW53241;
DT 03-AUG-1998 (first entry)
XX Homo sapiens vascular endothelial growth factor D (VEGF-D).
XX KM vascular endothelial growth factor; VEGF-D; angiogenesis;
XX KM modification; acceleration; wound healing; tissue; organ;
XX KM transplants; collateral circulation; infarction; arterial stenosis;
XX KM coronary artery disease; inhibition; cancer; treatment;
XX KM diabetic retinopathy; lung disorders; blood circulation;
XX KM gaseous exchange; chronic obstructive airway disease;
XX KM intestinal malabsorptive syndrome; biopsy; metastatic risk;
XX KM detection; diagnosis; congestive heart failure.
XX OS Homo sapiens.
XX PN MO9807832-A1.
XX PD 26-FEB-1998.
XX PE 21-AUG-1997; 97MO-US14696.
XX PR 01-JUL-1997; 97US-0051426.
XX PR 23-AUG-1996; 96AU-0001825.
XX PR 23-AUG-1996; 96US-0023751.
XX PR 11-NOV-1996; 96AU-0003554.
XX PR 14-NOV-1996; 96US-0031097.
XX PR 05-FEB-1997; 97AU-0004954.
XX PR 10-FEB-1997; 97US-0038814.
XX PR 19-JUN-1997; 97AU-0007435.
XX PA (LUDWIG) LUDWIG INST CANCER RES.
XX PA (UYHE-) UNIV HELSINKI LICENSING LTD.
XX PI Achen MG, Alltalo K, Stackner SA, Wilks AF;
XX DR WPI: 1998-179057/16.
XX DR N-PSDB; AAV20807.
XX PT New isolated vascular endothelial growth factor-D - used to develop
XX PT products for use in e.g. modifying angiogenesis or treating lung,
XX PT heart or intestinal disorders
XX PS Claim 16; Pages 60-61; 101pp; English.
XX CC The sequence is that of human lung vascular endothelial growth factor
XX CC D (VEGF-D). VEGF-D can be used for e.g. acceleration of angiogenesis
XX CC in wound healing, tissue or organ transplantation, or to establish
XX CC collateral circulation in tissue infarction or arterial stenosis,
XX CC such as coronary artery disease, and inhibition of angiogenesis in
XX CC the treatment of cancer or of diabetic retinopathy. It can also be
XX CC used in the treatment of lung disorders to improve blood circulation
XX CC in the lung and/or gaseous exchange between the lungs and O2 gas
XX CC permeability in cases of cardiac insufficiency, to improve blood
XX CC flow and gaseous exchange in chronic obstructive airway disease,
XX CC or to treat malabsorptive syndromes in the intestinal tract.
XX CC Quantitation of VEGF-D in cancer biopsy specimens may be useful
XX CC as an indicator of future metastatic risk. Antagonists can be used
XX CC for treating e.g. conditions such as congestive heart failure,
XX CC involving accumulations of fluid in the lung resulting from

```


CC Increases in vascular permeability. The products can also be used
CC for detection and diagnosis.

SO Sequence 354 AA:

Query Match 100.0%; Score 598; DB 19; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.4e-57;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYDIETLKVIDEWMORTQCSPRETCVEVASSELGKSTNTEFFKPCVNVFRCGGCCNEESLI 60
|||||
Db 93 fydieltkvideewqtrgtcspretcvevaseljgkstntffkpcvnmvfrcggcneesi 152

OY 61 CMTSTSYISKOLFETISVPLTSVPELVKPVVANHGTGCKLPAPRHPYS 109
|||||
Db 153 cmtstsyiskqlfetsvpltsvpelvpkvvanhtgcklptaprhyps 201

RESULT 7

AAW44293
ID AAW44293 standard; Protein; 354 AA.

AC AAW44293;

DT 22-JUN-1998 (first entry)

DE Human vascular endothelial growth factor D.

KW Human; vascular endothelial growth factor D; VEGF-D; gene therapy;

KM Inflammation; oedema.

OS Homo sapiens.

PN MO9802543-A1.

PD 22-JAN-1998.

PF 15-JUL-1997; 97MO-JP02456.

PR 15-JUL-1996; 96JP-0185216.

PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.

PI Hirata Y, Nezu J;

DR WPI; 1998-110591/10.

DR N-PSDB; AAV15156.

PT VEGF-D protein encoded by DNA - useful for, e.g. gene therapy and

PS Claim 1; Page 18-20; 52pp; Japanese.

CC The present sequence represents human vascular endothelial growth factor

CC D (VEGF-D). The VEGF-D protein, compounds and antibodies, which can bind

CC the protein, may be useful in, e.g. gene therapy and in treatment of

CC inflammation and oedema. Vectors, containing the VEGF-D DNA, and VEGF-D

CC DNA sequences may be used for screening for the compounds which bind to

CC the VEGF-D protein.

XX the VEGF-D protein.

SO Sequence 354 AA:

Query Match 100.0%; Score 598; DB 19; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.4e-57;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYDIETLKVIDEWMORTQCSPRETCVEVASSELGKSTNTEFFKPCVNVFRCGGCCNEESLI 60
|||||
Db 93 fydieltkvideewqtrgtcspretcvevaseljgkstntffkpcvnmvfrcggcneesi 152

OY 61 CMTSTSYISKOLFETISVPLTSVPELVKPVVANHGTGCKLPAPRHPYS 109

Db 153 cmtstsyiskqlfetsvpltsvpelvpkvvanhtgcklptaprhyps 201

RESULT 8

AAAB10649
ID AAAB10649 standard; Protein; 354 AA.

AC AAAB10649;

DT 19-JAN-2001 (first entry)

DE Human VEGD protein.

KW VEGF-X; vascular endothelial growth factor; human; vulnery; cyrostatic;

KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;

KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;

KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;

KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;

KW venous sore; diabetic ulcer; burns; skin graft growth; VEGD.

OS Homo sapiens.

PN WO200037641-A2.

PD 29-JUN-2000.

PF 21-DEC-1999; 99MO-US30503.

PR 22-DEC-1998; 98GB-0028377.

PR 18-MAR-1999; 99US-0124967.

PR 08-NOV-1999; 99US-0164131.

PA (JANC) JANSSEN PHARM NV.

PI Gordon RD, Sprengel JJ, Yon JR, Dijkmans JTH, Goslowska A;

PI Dhanaraj SN, Xu J;

DR WPI; 2000-442669/38.

PT New vascular endothelial growth factor protein, useful for treating or

PT preventing diseases associated with inappropriate angiogenesis activity

PS such as cancer, rheumatoid arthritis, psoriasis and wounds -

PS Disclosure; Fig 11; 127pp; English.

CC This invention describes a novel vascular endothelial growth factor-X

CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has

CC vulnery, cyrostatic, antirheumatic, antiarthritic, antipsoriatic and

CC antidiabetic activity and acts as an angiogenesis and vascularization

CC regulator. An antisense molecule of the invention is useful for treating

CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic

CC retinopathy by inhibiting angiogenic activity or inappropriate

CC vascularization including formation and proliferation of new blood

CC vessels, growth and development of tissues, tissue regeneration and organ

CC and tissue repair in a subject. The products of the invention are useful

CC for preparing medicaments for treating wounds such as dermal ulcers,

CC pressure sores, venous sores, diabetic ulcers and burns and to promote

CC skin graft growth, tissue repair, proliferation of new blood vessels,

CC tissue regeneration and organ repair by promoting angiogenic activity or

CC vascularization. This sequence represents the human VEGD protein used

CC to illustrate the method of the invention.

XX the VEGD protein.

SO Sequence 354 AA:

Query Match 100.0%; Score 598; DB 21; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.4e-57;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYDIETLKVIDEWMORTQCSPRETCVEVASSELGKSTNTEFFKPCVNVFRCGGCCNEESLI 60
|||||
Db 93 fydieltkvideewqtrgtcspretcvevaseljgkstntffkpcvnmvfrcggcneesi 152


```

Db      153 cmtstsyiskqlfetsvpltsvpelvpvkvanhgtckclptaripryps 201
|||||
RESULT 11
AA70983
XX ID AAY70983 standard; Protein; 354 AA.
AC AAY70983;
XX
XX 09-AUG-2000 (first entry)
DT
XX
XX Human vascular endothelial growth factor (VEGF)-D protein.
XX
XX Vascular endothelial growth factor-D; VEGF; human; re-endothelialisation;
XX vascular endothelial growth factor receptor; VEGFR; vascular trauma;
XX blood vessel; cardiovascular surgery; anti-restenosis agent; prevention;
XX restenosis; stenosis; percutaneous transluminal coronary angioplasty.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..21
XX /label= Signal_peptide
XX /note= "Cleavage results in partially-processed VEGF-D
XX protein"
XX 22..92
XX /label= Amino_terminal_peptide
XX /note= "Cleavage results in fully-processed mature
XX VEGF-D protein"
XX 93..201
XX /label= Recombinantly_matured_VEGF_D_protein
XX /note= "Processed vascular epithelial growth factor-D"
XX 202..354
XX /label= Carboxy_terminal_peptide
XX /note= "Cleavage results in partially-processed VEGF-D
XX protein"
XX
XX MO200024412-A2.
XX
XX 04-MAY-2000.
XX
XX 26-OCT-1999; 99MO-US24054.
XX
XX 26-OCT-1998; 98US-0105587.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX (UYHE-) UNIV HELSINKI LICENSING LTD OY.
XX (YLAH/) YLA-HERTUALA S.
XX
XX Yla-heretuala S, Ailtalo K, Hiltunen MO, Jeltsch MM, Achen MG;
XX WPI; 2000-350584/30.
XX DR N-PSDB; AAD00340.
XX
XX Preventing stenosis and restenosis in mammals using vascular
XX endothelial growth factor proteins or the nucleic acids encoding them -
XX
XX Disclosure; Page 53-55; 61pp; English.
XX
XX The present amino acid sequence is the complete human prepro-vascular
XX endothelial growth factor (VEGF)-D. VEGF-D has the ability to stimulate
XX re-endothelialisation of an injured blood vessel, without significant
XX stimulation of smooth muscle cell proliferation. It can bind to and
XX stimulate VEGFR-2 (vascular endothelial growth factor receptor) and/or
XX VEGFR-3 phosphorylation in cells that express such receptors. An
XX anti-restenosis agent comprising either a VEGF-D gene or protein is
XX used in a method to reduce or prevent restenosis and stenosis of a blood
XX vessel following vascular trauma e.g., cardiovascular surgery and
XX percutaneous transluminal coronary angioplasty.
XX
XX Sequence 354 AA:

```

```

Query Match 100.0%; Score 598; DB 21; Length 354;
Best Local Similarity 100.0%; Pred No. 1.4e-57;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDIETLKVIDEEMORTQCSPRETCVEVASELCKSTWTFKPCVNVFRGCCNESSL 60
Db 93 fydietlkvideemqrtqcspretcvevaseelgksntlffkpcvnmvfrggsnesll 152
QY 61 CMMTSTSYISKQLFETSVPLTSVPVLPVKNVANHGTCKCLPTARIPRYP 109
Db 153 cmtstsyiskqlfetsvpltsvpelvpvkvanhgtckclptaripryps 201

RESULT 12
AA797573
XX ID AAY97573 standard; Protein; 354 AA.
AC AAY97573;
XX
XX 05-APR-2001 (first entry)
DT
XX
XX Human VEGF-D1 protein sequence.
XX
XX Human; angiogenic protein; wound healing; vascular tissue repair;
XX peripheral arterial disease; critical limb ischaemia; coronary disease;
XX angiogenesis; tumour; inflammation; diabetic retinopathy; psoriasis;
XX rheumatoid arthritis; autoimmune disease; allergy; cancer; therapy;
XX infectious disease; neurodegeneration;
XX vascular endothelial growth factor-D1; VEGF-D1.
XX
XX Homo sapiens.
XX
XX MO200075163-A1.
XX
XX 14-DEC-2000.
XX
XX 01-JUN-2000; 2000MO-US14925.
XX
XX 03-JUN-1999; 99US-0137796.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Hu J, Cao L;
XX WPI; 2001-071057/08.
XX DR N-PSDB; AAA91007.
XX
XX New nucleic acid encoding angiogenic proteins, useful e.g. for
XX promoting healing of wounds and treating peripheral arterial disease,
XX critical limb ischaemia or coronary disease -
XX
XX Claim 11; Page 228-229; 244pp; English.
XX
XX This sequence is vascular endothelial growth factor-D1 (VEGF-D1),
XX which is an angiogenic protein of the invention. The angiogenic proteins
XX and the DNA sequences encoding them, are used to prevent, treat or
XX ameliorate disease and to detect diseases, or susceptibility, by
XX detecting mutations or the presence or amount of angiogenic protein
XX expression. Particularly they are used to stimulate wound healing,
XX growth of damaged bone and tissue, and for repair of vascular tissue,
XX especially peripheral arterial disease, critical limb ischaemia or
XX coronary disease. Antagonists of the sequences are used to inhibit
XX angiogenesis in tumours and to treat inflammation (where associated with
XX increased vascular permeability), diabetic retinopathy, rheumatoid
XX arthritis or psoriasis. Agonists are also useful for stimulating
XX lymphangiogenesis. The proteins are also used to identify specific
XX binding agents (potential therapeutic agents) and to raise antibodies.
XX The antibodies are useful as therapeutic (antagonists; for detection,
XX purification and targeting of proteins in vivo or in vitro diagnosis
XX (including imaging) or for therapy (including when linked to e.g. a label
XX or cytotoxin); and for immunotyping of cells, e.g. for detecting minimal
XX residual disease or haematopoietic progenitor/stem cells. It is also

```


SQ Sequence 354 AA:

Query Match 97.3%; Score 582; DB 22; Length 354;
Best Local Similarity 97.2%; Pred. No. 8.1e-56;
Matches 106; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 FVDIETLKVDEEMORTQCSPRETCVEVASSELGKSTNTEFFKPPCVNFRGCCNEESLI 60
|||||
Db 93 fyietlkvideewqrtgcspretcvaseisgkstntffkppcvnvercgcneesfm 152
|||||

OY 61 CMTSTSYSKOLFELISVPLTSPVELVPVKVANHGCCKLPNAPRHPYS 109
|||||
Db 153 cmtstsyiskqlfelsvpltsvpelvpykvanhgtckclptaprhyps 201
|||||

RESULT 15
AAW14994
ID AAW14994 standard; Protein; 620 AA.
XX
AC AAW14994;
XX
DT 05-JUL-1997 (first entry)
XX
DE Human c-Fos induced growth factor (clone HF175 ORF2 product).
XX
XX c-Fos induced growth factor; FIGF; Fos regulated gene;
KM proto-oncogene; lung disorder; cancer; tumour; therapy;
KM antibody; Transgenic animal.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 16 /note= "residue 16 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 26 /note= "residue 26 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 29 /note= "residue 29 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 47 /note= "residue 47 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 71 /note= "residue 71 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 72 /note= "residue 72 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 76 /note= "residue 76 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 136 /note= "residue 136 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 220 /note= "residue 220 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 341 /note= "residue 341 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 344 /note= "residue 344 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 377 /note= "residue 377 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 435 /note= "residue 435 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 486 /note= "residue 486 corresponds to an in-frame
stop codon in reading frame 2 of HF175"

FT /note= "residue 486 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 497 /note= "residue 497 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 518 /note= "residue 518 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 541 /note= "residue 541 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 553 /note= "residue 553 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 557 /note= "residue 557 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 562 /note= "residue 562 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 579 /note= "residue 579 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 592 /note= "residue 592 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 593 /note= "residue 593 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 597 /note= "residue 597 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 605 /note= "residue 605 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
FT Misc-difference 608 /note= "residue 608 corresponds to an in-frame
stop codon in reading frame 2 of HF175"
PN WO9712972-A2.
XX 10-APR-1997.
PD 30-SEP-1996; 96WC-1B01113.
XX 13-JUN-1996; 96GB-0012368.
PR 29-SEP-1995; 95GB-0019928.
PA (UYSI-) UNIV SIENA.
XX
PI Oliviero S;
XX
XX WPI: 1997-226216/20.
DR N-PSDB; AAT62961.
PT Nucleotide molecule encoding c-Fos induced growth factor protein -
useful in therapy; in manufacture of compositions for treatment of
developmental disorders and in generation of transgenic animal
XX
XX Claim 3; Fig 2; 64pp; English.
CC 3 Polypeptide sequences (AAW14993-95) are the respective translated
sequences of reading frames 1, 2 and 3 of clone HF175 (AAT62961), the
human homologue of murine clone F0401 (AAT62960), which codes for a
novel c-Fos induced growth factor (FIGF) (see also AAW14992).
CC Examination of the 3 polypeptides indicates that reading frame 2
has the longest region free of nonsense codons. FIGF is a c-fos-
dependent autocrine growth factor able to induce cell division
entry and, when over-expressed, a transformed phenotype in
fibroblasts. It could be implicated in tumours and development.
CC Recombinant FIGF can be produced in transformed host (e.g. CHO
cells). It can be used to identify its receptors and in an assay
for the identification of agonists and antagonists. Antibodies

CC raised against FIGF can be used to block the function of the
CC protein and thereby inhibit or suppress tumour growth. Transgenic
CC animals expressing FIGF can be generated for use e.g. as models for
CX research.

sq Sequence 620 AA;

Query Match	97.3%;	Score 582;	DB 18;	Length 620;
Best Local Similarity	97.2%;	Pred. NO. 1.6e-55;		
Matches 106;	Conservative 1;	Mismatches 2;	Indels 0;	Gaps 0;

```
OY 1 FYDIETLEKVIDEEMORQOSPRECECVASASTLGKSTTFKPRPCVANNRFGCGCNEESI 60
Db 173 fydieltlvideeqrtqspreccevasaelystntffkpcvnnrercgscneesi 222
OY 61 CMNSTSVTSKOLFELISVPLTSVPELTPVAVANTGCKCPTAPRHYS 109
Db 233 cmnstsvtskqlfelisvpltsvpeelpvkvavantgckcpcpbrphtys 281
```

Search completed: October 17, 2001, 14:46:02
Job time: 128 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:47:11 ; Search time 62.93 seconds
(without alignments)
35.664 Million cell updates/sec

Title: US-09-427-657-4_COPY_93_201

Perfect score: 598
Sequence: 1 FYDIEFLKVIDEEMQRTGCS.....KVAHHTGCKCLPTAPRHRYS 109

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pdp:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pdp:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pdp:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pdp:*
5: /cgn2_6/ptodata/2/1aa/PTCUS.COMB.pdp:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pdp:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	598	100.0	325	4	US-08-915-795-3 Sequence 3, Appl 1
2	598	100.0	354	4	US-08-915-795-5 Sequence 5, Appl 1
3	575	96.2	321	4	US-08-915-795-9 Sequence 9, Appl 1
4	575	96.2	358	2	US-08-915-795-8 Sequence 8, Appl 1
5	377	63.0	350	2	US-08-999-811-4 Sequence 2, Appl 1
6	377	63.0	350	2	US-08-824-996-2 Sequence 4, Appl 1
7	377	63.0	350	3	US-09-042-105-4 Sequence 4, Appl 1
8	377	63.0	350	4	US-08-510-133A-33 Sequence 3, Appl 1
9	377	63.0	350	4	US-08-585-895-33 Sequence 3, Appl 1
10	377	63.0	419	2	US-08-999-811-2 Sequence 2, Appl 1
11	377	63.0	419	3	US-09-042-105-2 Sequence 2, Appl 1
12	377	63.0	419	3	US-09-042-105-18 Sequence 18, Appl 1
13	377	63.0	419	4	US-08-795-430-8 Sequence 8, Appl 1
14	377	63.0	419	4	US-08-510-133A-35 Sequence 35, Appl 1
15	377	63.0	419	5	PCT-US96-09001-2 Sequence 2, Appl 1
16	374	62.5	415	4	US-08-795-430-11 Sequence 11, Appl 1
17	355	59.4	418	4	US-08-795-430-13 Sequence 13, Appl 1
18	171.5	28.7	146	4	US-08-586-039B-33 Sequence 33, Appl 1
19	171.5	28.7	190	4	US-08-586-039B-31 Sequence 31, Appl 1
20	171.5	28.7	214	4	US-08-586-039B-35 Sequence 35, Appl 1
21	169.5	28.3	121	6	5194596-19 Patent No. 5194596
22	169.5	28.3	121	6	5219739-20 Patent No. 5219739
23	169.5	28.3	145	3	US-08-784-551C-2 Sequence 2, Appl 1
24	169.5	28.3	147	3	US-08-807-992B-1 Sequence 1, Appl 1
25	169.5	28.3	165	6	5194596-18 Patent No. 5194596
26	169.5	28.3	165	6	5219739-19 Patent No. 5219739
27	169.5	28.3	191	3	US-08-567-200A-2 Sequence 2, Appl 1

28	169.5	28.3	191	3	US-08-807-992B-2 Sequence 2, Appl 1
29	169.5	28.3	191	4	US-08-691-794-2 Sequence 2, Appl 1
30	169.5	28.3	191	4	US-08-795-430-56 Sequence 56, Appl 1
31	169.5	28.3	191	6	5332671-4 Patent No. 5332671
32	169.5	28.3	214	6	5240848-11 Patent No. 5240848
33	169.5	28.3	215	3	US-08-807-992B-3 Sequence 3, Appl 1
34	169.5	28.3	215	4	US-08-586-039B-49 Sequence 49, Appl 1
35	169.5	28.3	215	6	5219739-22 Patent No. 5219739
36	169.5	28.3	215	6	5240848-7 Patent No. 5240848
37	169.5	28.3	231	5	PCT-US96-09001-10 Sequence 10, Appl 1
38	169.5	28.3	232	2	US-08-999-811-7 Sequence 7, Appl 1
39	169.5	28.3	232	2	US-08-824-996-9 Sequence 9, Appl 1
40	169.5	28.3	232	3	US-08-807-992B-4 Sequence 4, Appl 1
41	169.5	28.3	232	3	US-09-042-105-7 Sequence 7, Appl 1
42	168	28.1	189	1	US-08-469-427A-15 Sequence 15, Appl 1
43	167.5	28.0	109	3	US-08-691-794-3 Sequence 3, Appl 1
44	166.5	27.8	120	6	5194596-9 Patent No. 5194596
45	166.5	27.8	120	6	5219739-9 Patent No. 5219739

ALIGNMENTS

RESULT 1
US-08-915-795-3
; Sequence 3, Application US/08915795
; Patent No. 6235713
; GENERAL INFORMATION:
; APPLICANT: Marc G. ACHEN
; APPLICANT: Andrew F. WILKS
; APPLICANT: Steven A. STACKER
; APPLICANT: Karl ALITALO
; TITLE OF INVENTION: GROWTH FACTOR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
; STREET: 1200 G Street, NW, Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: United States of America
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,795
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 1064/42983
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: Human Breast
; US-08-915-795-3
Query Match 100.0%; Score 598; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 1.3e-62;


```

? APPLICATION NUMBER: US 08/207,550
? FILING DATE: 8-MAR-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/465,968
? FILING DATE: 06-JUN-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: MARKOWICZ, KAREN R.
? REGISTRATION NUMBER: 36,351
? REFERENCE/DOCKET NUMBER: 1488.1000004
TELECOMMUNICATION INFORMATION:
    TELEPHONE: (202)371-2600
    TELEFAX: (202)371-2540
? INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
? MOLECULE TYPE: protein
us-08-999-811-4

Query Match          63.0%; Score 377; DB 2; Length 350;
Best Local Similarity 63.6%, Pred. No. 1,4e-36;
Matches 63; Conservative 15; Mismatches 21; Indels 0; Gaps 0;

OY      2 YDIELKLVIDEWMORTQCSPRETCVENVSELGKSITNFFPKPCPVNVFPKGCCNEESTLIC 61
        | : |||:||||::|||:||||:|:| |:|:|||||||::|||:||||| | | 
Db       45 INTELKSIDENMRKTGCMPEVCIDYGKEGVATNTTFRKPPCSVYRCGGCNSBGLQC   104
        |||||::|||:|||||:|:| |:|:|||||||::|||:||||| | | 

OY      62 MNISTSYISKQLFEISLVPETSLVPDLVPIKYVANHTGCCKL   100
        |||||::|||:|||||:|:~|:|:|||||||::|||:||||| | | 
```

```

Query March 63.0% Score 377; DB 2; Length 350;
Best Local Similarity 63.6% Pred. No. 1,4e-36;
Matches 63; Conservative 15; Mismatches 21; Indels 0; Gaps 0;

OY 2 YDLEFLKVIDEEMORTQCSPRETCVEVASLEGKSTNTPFKPCYVNFRCGCCNEESTLIC 61
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 45 YNTLTKSIDENMRKTCMPREVCIDVKGEGKGVATNTFFKPCVSVYRCGCCNSEGLQC 104
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 62 MNTSTSYSKOLFETSYVPLTSPVELVPVKVANHNGCKL 100
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 105 MNTSTSYSLKTLFETLTVPLDSGPKPVTISFANHNSCRQM 143
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 6
US-08-824-996-2
; Sequence 2, Application US/088249968
; Patent No. 5935820
; GENERAL INFORMATION:
; APPLICANT: Hu, Jing-Shan
; APPLICANT: Rosen, Craig A.
; APPLICANT: Cao, Liang
; TITLE OF INVENTION: Polynucleotides Encoding Vascular Endothelial Growth
; TITLE OF INVENTION: Factor 2

```

```

; SEQ ID NO 2
; LENGTH: 350
; TYPE: PR1
; ORGANISM: Homo sapiens
US-08-824-996-2

Query Match      63.0%; Score 377; DB 2; Length 350;
Best Local Similarity 63.6%; Pred. No. 1,4e-36;
Matches 63; Conservative 15; Mismatches 21; Indels 0; Gaps 0;

QY      2 YDIFTLKVIDEEMORTCCSPRETCEVAASELKGSTNFFPKPPCVNVPFGGCCNEESLIG 61
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      45 YNTFLKSLIDEMWRTQCCMPREVCIDVKGKEGVATNFFPKPPCVSVYRCGGCCNSELQ 104

QY      62 MNTSTSYISKOLEISVPLTSPPELVYVKVANHGGCCCL 100
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      105 MNTSTSYLSKTLFEITVPLSGPKPVTISFANHSTSCROM 143

RESULT      7
US-09-042-105-4

```



```

1  ATTORNEY/AGENT INFORMATION:
2  NAME: Gass, David A.
3  REGISTRATION NUMBER: 38,153
4  REFERENCE/DOCKET NUMBER: 28113/33072
5  TELECOMMUNICATION INFORMATION:
6  TELEPHONE: 312/474-6300
7  TELEFAX: 312/474-0448
8  TELEX: 25-3856
9  INFORMATION FOR SEQ. ID NO.: 33:
10 SEQUENCE CHARACTERISTICS:
11     LENGTH: 350 amino acids
12     TYPE: amino acid
13     TOPOLOGY: linear
14     MOLECULE TYPE: protein
15     DS-08-585-895-33

```

Query Match	63.0%	Score 377	DB 4	Length 350
Best Local Similarity	63.6%	Pred. No. 1.4e-36		
Matches	63	Conservative	15	Mismatches 21
			Indels	0
			Gaps	0

Qy 2 YDIELKIKYIDEMOQRTSPETCEVAASELGKTNFFPPCVNVRGCCMEELIC 61
Db 105 YNTELLKIDNEMKRTQCMPEVCIIDVCKEGCATNFFRPCTSVYRCGCCNSELQC 104

RESULT 10
US-08-999-811-2
; Sequence 2, Application US/08999811

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

```

;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
US-08-999-811-2

```

Query Match	63.0%;	Score 377;	DB 2;	Length 419;
Best Local Similarity	63.6%;	Pred. No. 1.8e-36;		
Matches	63;	Conservative 15;	Mismatches 21;	Indels 0;
			Gaps	0;

[illegible]

RESULT 11
US-09-042-105-2、
; Sequence 2, Application US/09042105
; Patent No. 6040157

APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:

STATE: DC
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.33
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/042,105
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550
FILING DATE: 8-MAR-1994
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995

FILING DATE: 06-JUN-1993
 CLASSIFICATION:
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: TO BE ASSIGNED
FILING DATE: 24-DEC-1997
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: ERIC K. STEFFE

REGISTRATION NUMBER: 36,888
REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEO ID NO: 3

SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-09-042-105-2

Query Match 63.08; Score 377; DB 3; Length 419;

Db 114 YNTEILKSIDNEMKRTQMPREVCIDVGKEGVATNFFRPKPCSVYRGCGCCNSELQC 173

Qy 62 MNTSTSYISKQLEISVPLVPELVKPYKANHTGCKL 100
|||||:| | | | | | | | : | : | | | | : | :
174 MNTSTSYLSKTLFELITVPLSGRPVITISFANHTSCGM 212

```

RESULT 14
US-08-510-133A-35
Sequence 35, Application US/08510133A
Patent No. 6221839
GENERAL INFORMATION:
APPLICANT: Altalo, Karl
Joukov, Vladimir
TITLE OF INVENTION: Receptor Ligand
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshail, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
City: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/510,133A
FILING DATE: 01-Aug-1995
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28113/32863
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-08-510-133A-35

```

	Query Match	63.0%;	Score 377;	DB 4;	Length 419;	
	Best Local Similarity	63.6%;	Pred.No. 1.8e-36;			
Matches	63;	Conservative	15;	Mismatches	21;	Indels 0; Gaps 0
OY	2	YDIELKXVIDEWMOTQCSPRETCVVASLSEKSNTPFKPCVNVNFRGGCGNSESLIC	61			
Dd	114	YNTELKSIDNDMKRTQCMPEVCIDVGKEFVAIVNTFPKPPCVSVYTRGCGCNSBGLQC	173			
OY	62	MNTSTSYISKOLEIISVPILTSVPELVPAKVANHTGCCKL	100			
Dd	174	MNTSTSYISKLTLEITVPLSQPKPVTLSFAHNHSCRCM	212			

```

RESULT 15
PCT-US96-09001-2
: Sequence 2, Application PC/TUS9609001
: GENERAL INFORMATION:
: APPLICANT: HU, ET AL.
: TITLE OF INVENTION: Human Vascular EndothelialGrowth Factor 2
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

```

```

1 ADDRESSSEE: CECCHI, STEWART & OLSTEN
2 STREET: 6 BECKER FARM ROAD
3 CITY: ROSELAND
4 STATE: NEW JERSEY
5 COUNTRY: USA
6 ZIP: 07068
7
8 COMPUTER READABLE FORM:
9 MEDIUM TYPE: 3.5 INCH DISKETTE
10
11 COMPUTER: IBM PS/2
12 OPERATING SYSTEM: MS-DOS
13 SOFTWARE: WORD PERFECT 5.1
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER: PCT/US96/09001
16 FILING DATE:
17 CLASSIFICATION:
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER: 08/465,968
20 FILING DATE: 6 JUN 95
21 APPLICATION NUMBER: 08/207,550
22 FILING DATE: 8 MAR 1994
23 ATTORNEY/AGENT INFORMATION:
24 NAME: FERRARO, GREGORY D.
25 REGISTRATION NUMBER: 36,134
26 REFERENCE/DOCKET NUMBER: 325800-288
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: 201-994-1700
29 TELEFAX: 201-994-1744
30 INFORMATION FOR SEQ ID NO: 2:
31 SEQUENCE CHARACTERISTICS:
32 LENGTH: 419 AMINO ACIDS
33 TYPE: AMINO ACID
34 STRANDEDNESS:
35 TOPOLOGY: LINEAR
36 MOLECULE TYPE: PROTEIN
37
38 PCT-US96-09001-2

```

[illegible]

Search completed: October 17, 2001, 14:47:12
Job time: 198 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:48:42 ; Search time 78.16 Seconds

(Without alignments)
106.231 Million cell updates/sec

Title: US-09-427-657-4_COPY_93_201

Perfect score: 598

Sequence: 1 FVDIEFLKVIDEMQRTQCS.....KVAHNGCKCLPTARHPHS 109

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	377	63.0	419	2	S69207
2	173.5	29.0	190	2	B44881
3	173.5	29.0	214	2	A44881
4	171.5	28.7	190	2	A35987
5	169.5	28.3	232	2	A41551
6	167.5	28.0	146	2	S57956
7	166.5	27.8	130	2	A33787
8	166.5	27.8	190	2	B40080
9	164	27.4	148	2	D49530
10	163.5	27.3	190	2	S52130
11	160	26.8	149	2	A4136
12	147.5	24.7	188	2	JC4680
13	147.5	24.7	207	2	JC4679
14	145.5	24.3	158	2	A56125
15	144.5	24.2	133	2	B49530
16	124.5	20.3	241	1	PFHUG2
17	121.5	20.3	245	1	TVCTSS
18	119.5	20.0	226	1	TVNVS
19	113.5	19.0	225	1	S25097
20	112.5	18.8	196	2	B28964
21	112.5	18.8	211	1	PFHUG1
22	112.5	18.8	241	1	PFHUGB
23	110.5	18.5	166	2	JN0248
24	110.5	18.5	186	2	JN0735
25	108.5	18.1	271	2	A25669
26	106.5	17.8	197	2	S25096
27	105.5	17.5	196	2	A48851
28	104.5	17.5	196	2	A37359
29	104.5	17.5	200	2	T51551

30	104.5	17.5	215	2	S08220	platelet-derived g
31	104.5	17.5	226	2	T51550	platelet-derived g
32	100	16.7	128	2	T51295	vascular endotheli
33	98	16.4	120	2	A39555	glycoprotein hormo
34	89	14.9	93	2	S07091	glycoprotein hormo
35	81	13.5	117	2	A37198	glycoprotein hormo
36	80.5	13.5	96	2	S74086	foliitropin alpha
37	80.5	13.5	120	2	A45585	glycoprotein hormo
38	80.5	13.5	120	2	T51241	pituitary glycopro
39	80	13.4	117	2	T50992	gonadotropin alpha
40	79.5	13.3	116	1	TTH0AP	glycoprotein hormo
41	78.5	13.1	118	2	S16762	gonadotropin alpha
42	78.5	13.1	118	2	A60626	glycoprotein hormo
43	77	12.9	372	2	T23680	hypothetical prote
44	76	12.7	118	1	UTCRA	glycoprotein hormo
45	76	12.7	118	2	A40554	glycoprotein hormo

ALIGNMENTS

```

RESULT 1
S69207
vascular endothelial growth factor C precursor - human
N:Alternate names: FLT4 ligand DHM
C:Species: Homo sapiens (man)
C>Date: 27-Apr-1996 #sequence,revision 01-Nov-1996 #text,change 08-Oct-1999
C:Accession: S69207; S61795; S71443; S69208; G02659
R:Joukov, V.; Pajusola, K.; Kaipainen, A.; Chillov, D.; Lahtinen, I.; Kukk, E.; Saksel
EMBO J. 15, 1751, 1996
A>Title: A novel vascular endothelial growth factor, VEGF-C, is a ligand
A:Reference number: S69207; MUID:96203094
A:Accession: S69207
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-419 <JOU>
A:Cross-references: EMBL:X94216; NID:g1177488; PIDN:CAA63907.1; PID:e221096; PID:g118
A>Note: this is a revision to the sequence from reference S61795
R:Joukov, V.; Pajusola, K.; Kaipainen, A.; Chillov, D.; Lahtinen, I.; Kukk, E.; Saksel
EMBO J. 15, 290-298, 1996
A>Title: A novel vascular endothelial growth factor, VEGF-C, is a ligand for the FLT4
A:Reference number: S61795; MUID:96178224
A:Accession: S61795
A>Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 70-419 <JOU>
A:Cross-references: EMBL:70-419 <JOU>
A>Note: this sequence has been revised in reference S69207
A:Accession: S71443
A:Molecule type: protein
A:Residues: X',104-120 <JOU>
R:Lee, J.; Gray, A.; Yuan, J.; Luch, S.M.; Avraham, H.; Wood, W.I.
submitted to the EMBL Data Library, December 1995
A>Description: Vascular endothelial growth factor related protein (VRP): A ligand and
A:Reference number: S69208
A:Accession: S69208
A:Molecule type: mRNA
A:Residues: 1-419 <LEE>
A:Cross-references: EMBL:043142; NID:g1150988; PIDN:AAA85214.1; PID:g1150989
R:Morris, J.C.
submitted to the EMBL Data Library, May 1996
A:Reference number: H01557
A:Accession: G02659
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-419 <MOR>
A:Cross-references: EMBL:U58111; NID:g1373426; PIDN:AAB02909.1; PID:g1373427
C:Genetics:
A:Gene: GDB:VEGFC; VRP
A:Cross-references: GDB:3890883; OMIM:601528
F:1-12/Domain: signal sequence #status predicted <SIG>
F:13-102/Domain: propeptide #status predicted <PRO>

```


Db 97 SESNVTMÖIMRIK-PHQ5-QHIGEMSFLOHNSRECRPKKDR 135

RESULT 5

AA1551
vascular endothelial growth factor 206 precursor - human
N:Alternate names: vascular permeability factor
M:Controls: vascular endothelial growth factor 121 (VEGF 121); VEGF 165; VEGF 189; VEGF
C:Species: Homo sapiens (man)
C:Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 05-Nov-1999
A:Accession: AA1551; C41551; A40454; B40454; C40454; A40079; A40080; J01463; J01
R: Houck, K.A.; Ferrara, N.; Winer, J.; Cachianes, G.; Li, B.; Leung, D.W.
Mol. Endocrinol. 5, 1806-1814, 1991
A:Title: The vascular endothelial growth factor family: identification of a fourth molec
A:Reference number: AA1551; MUID:92168077
A:Accession: AA1551
A:Molecule type: mRNA
A:Residues: 1-232 <HOU1>
A:Cross-references: GB:S85192; NID:9246155; PID:9246156
A:Accession: C41551
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-140, 'N', 183-232 <HOU2>
A:Accession: B41551
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-141,227-232 <HOU>
R:Ritscher, E.; Mitchell, R.; Hartman, T.; Silva, M.; Gospodarowicz, D.; Fiddes, J.C.; At
J. Biol. Chem. 266, 11947-11954, 1991
A:Title: The human gene for vascular endothelial growth factor. Multiple protein forms a
A:Reference number: A40454; MUID:91268072
A:Accession: A40454
A:Molecule type: DNA
A:Residues: 1-165,183-232 <T11>
A:Cross-references: GB:M63972; GB:M63973; GB:M63974; GB:M63975; GB:M63976; GB
A:Accession: B40454
A:Molecule type: DNA
A:Residues: 1-140, 'N', 183-232 <T12>
A:Cross-references: GB:M63972; GB:M63973; GB:M63974; GB:M63975; GB:M63977; GB
A:Accession: C40454
A:Molecule type: DNA
A:Residues: 1-141,227-232 <T13>
A:Cross-references: GB:M63972; GB:M63973; GB:M63974; GB:M63975; GB:M63978
R:Kech, P.J.; Huser, S.D.; Krivi, G.; Sanzo, K.; Warren, T.; Feder, J.; Connolly, D.T.
Science 246, 1309-1312, 1989
A:Title: Vascular permeability factor, an endothelial cell mitogen related to PDGF.
A:Reference number: A40079; MUID:90069609
A:Accession: A40079
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-165,183-232 <KRC>
A:Cross-references: GB:M27281; NID:9340300; PIDN:AAA36807.1; PID:9340301
R:Leung, D.W.; Cachianes, G.; Kiang, W.-J.; Goeddel, D.V.; Ferrara, N.
Science 246, 1306-1309, 1989
A:Title: Vascular endothelial growth factor is a secreted angiogenic mitogen.
A:Reference number: A40080; MUID:90069608
A:Accession: A40080
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-140, 'N', 183-232 <LEU>
A:Cross-references: GB:M32977; NID:9181970; PIDN:AAA35789.1; PID:9181971
R:Weindel, K.; Marne, D.; Welch, H.A.
Biochem. Biophys. Res. Commun. 183, 1167-1174, 1992
A:Title: AIDS-associated Kaposi's sarcoma cells in culture express vascular endothelial
A:Reference number: J01463; MUID:92231879
A:Accession: J01463
A:Molecule type: mRNA
A:Residues: 1-140, 'N', 183-232 <ME1>
A:Cross-references: EMBL:X62568; NID:937658; PIDN:CAA44447.1; PID:937659
A:Experimental source: AIDS-Kaposi's sarcoma cell
A:Accession: J01464
A:Molecule type: mRNA

A:Residues: 1-140, 'N', 227-232 <ME2>

A:Experimental source: AIDS-Kaposi's sarcoma cell

R:Connolly, D.T.; Olander, J.V.; Heuvelman, D.; Nelson, R.; Monsell, R.; Siegel, N.;

J. Biol. Chem. 264, 20017-20024, 1989

A:Title: Human vascular permeability factor. Isolation from U937 cells.

A:Reference number: A34492; MUID:90062112

A:Accession: A34492

A:Molecule type: protein

A:Residues: 27-36,43-49, 'R', 72-76, 'Q', 78-81,59-71 <CON>

C:Comment: The most common of several alternatively spliced forms is VEGF 165.

C:Genetics:

A:Gene: GDB:VEGF

A:Cross-references: GDB:132244; OMIM:192240

A:Map position: 6p21-6p12

C:Function:

A:Description: promotes fluid and protein leakage from blood vessels

C:Keywords: alternative splicing; angiogenesis; dimer; disulfide bond; extracellular

F:1-232/Product: vascular endothelial growth factor 206 precursor #status predicted <

F:1-165,183-232/Product: vascular endothelial growth factor 189 precursor #status pre

F:1-141,227-232/Product: vascular endothelial growth factor 121 precursor #status pre

F:1-26/Domain: signal sequence #status predicted <SIG>

F:101/Binding site: carbonylrate (asn) (covalent) #status predicted

Query Match 28.3%; Score 169.5; DB 2; Length 232;

Best Local Similarity 35.2%; Pred. No. 66-10;

Matches 37; Conservative 14; Mismatches 43; Indels 11; Gaps 3;

QY 5 ETLKVIDEMQRTQCSPRETCVEVASELQSTNTFFKPPCVNVFRGCGCCNESLTCMNT 64

DB 39 EVKFMQD-VYQSFCHRIETLVDFQEPDEIRYIFKPCVPLMRGCGCCNDGLSCVPT 97

QY 65 STSYISKOLFELISVPLTSP----ELVPKVNANHTGCKLPTAPR 105

DB 98 EESNITMÖIMRIK-----PHQOHIGEMSFLOHNSRECRPKKDR 136

RESULT 6

S57956

ovine vascular endothelial growth factor - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 05-Nov-1999

C:Accession: S57956

R:Redner, D.A.; Dai, Y.; Li, J.; Jones, S.C.; Moor, R.M.

submitted to the EMBL Data Library, July 1995

A:Reference number: S57956

A:Accession: S57956

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-146 <RED>

A:Cross-references: EMBL:X89506; NID:9899350; PIDN:CAA61677.1; PID:9899351

Query Match 28.0%; Score 167.5; DB 2; Length 146;

Best Local Similarity 38.1%; Pred. No. 66-10;

Matches 37; Conservative 14; Mismatches 43; Indels 3; Gaps 3;

QY 5 ETLKVIDEMQRTQCSPRETCVEVASELQSTNTFFKPPCVNVFRGCGCCNESLTCMNT 64

DB 38 EVKFMQD-VYQSFCHRIETLVDFQEPDEIRYIFKPCVPLMRGCGCCNDGLSCVPT 96

QY 65 STSYISKOLFELISVPLTSPVELVPKVNANHTGCKLPTAPR 101

DB 97 EEPNITMÖIMRIK-PHQ5-QHIGEMSFLOHNSRECRPKKDR 131

RESULT 7

A33787

vascular endothelial growth factor (version 1) - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 16-Mar-1990 #sequence_revision 16-Mar-1990 #text_change 05-Nov-1999

C:Accession: A33787

R:Ritscher, E.; Gospodarowicz, D.; Mitchell, R.; Silva, M.; Schilling, J.; Lau, K.; Cr

A:Note: sequence extracted from NCBI backbone (NCBIN:141420, NCBIIP:141425)

```
Query Match      24.28; Score 144.5; DB 2; Length 133;
Best Local Similarity 34.18; Pred. No. 1.2e-07;
Matches 31; Conservative 14; Mismatches 41; Indels 5; Gaps 2;

OY 17 TCCSPRETCVAVASELGSSTNTFFKPPCVAVFRGCGCCNEESLTCMNTSTSYISKOLFEL 76
   :| | | | | : | | | | | | | | | | | | | | | | | | | | | | | |
Db 34 SECKPRPIVVPVSETHPELTSQRFNPPCYTLMKCGGCCNDSELCVPTLEVAVTMELGA 93
   :| | | | | : | | | | | | | | | | | | | | | | | | | | | | | |
OY 77 SVPLTSVPPELVPKVANHTGCKCLP--TAP 104
   | | | | | : | | | | | | | | | | | | | | | | | | | | | | | |
Db 94 SG--SGSNGMORLSFVEHKKKDCRPRFTTP 122
   | | | | | : | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: October 17, 2001, 14:48:42
Job time: 288 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:51:47 ; Search time 42.92 Seconds
(without alignments)
86.996 Million cell updates/sec

Title: US-09-427-657-4_COPY_93_201
Perfect score: 598
Sequence: 1 FYDIEFLKVIDEWMQRTQCS.....KVAHNGCKCLPTAPRHPYS 109

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 93435 seqs, 34255486 residues
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	377	63.0	419	1	VEGC_HUMAN
2	374	62.5	415	1	VEGC_MOUSE
3	173.5	29.0	214	1	VEGF_MOUSE
4	171.5	28.7	190	1	VEGF_RAT
5	169.5	28.3	215	1	VEGF_MOUSE
6	167.5	28.0	146	1	VEGF_SHEEP
7	166.5	27.8	190	1	VEGF_BOVIN
8	164	27.4	148	1	VEGF_ORF7
9	163.5	27.3	190	1	VEGF_PIG
10	159	26.6	170	1	PLGF_HUMAN
11	155.5	26.0	164	1	VEGF_CAVPO
12	147.5	24.7	133	1	VEGH_ORF2
13	147.5	24.7	188	1	VEGH_MOUSE
14	146.5	24.5	188	1	VEGH_HUMAN
15	140.5	23.5	158	1	PLGF_MOUSE
16	129.5	21.7	216	1	VEGF_CHICK
17	124.5	20.8	241	1	PDGF_HUMAN
18	121.5	20.3	245	1	PDGF_MOUSE
19	119.5	20.0	226	1	PDGF_SHEEP
20	116.5	19.5	225	1	PDGF_RAT
21	113.5	19.0	225	1	PDGF_MOUSE
22	112.5	18.8	211	1	PDGF_HUMAN
23	112.5	18.8	241	1	PDGF_MOUSE
24	110.5	18.5	213	1	PDGF_RABIT
25	107.5	18.0	211	1	PDGF_MOUSE
26	106.5	17.8	204	1	PDGF_RAT
27	104.5	16.5	126	1	PDGF_MOUSE
28	98	16.4	120	1	GLHA_MACW
29	89	14.9	93	1	GLHA_MOUSE
30	81	13.5	117	1	GLHA_ANGAN
31	80.5	13.5	96	1	GLHA_STRCA
32	80.5	13.5	120	1	GLHA_MELGA
33	79.5	13.3	116	1	GLHA_HUMAN

ALIGNMENTS

RESULT	1	STANDARD:	PRT:	419 AA.
VEGC_HUMAN				
ID	VEGC_HUMAN			
AC	P49767			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR (VEGF-C) (VASCULAR			
DE	ENDOTHELIAL GROWTH FACTOR RELATED PROTEIN) (VRP) (FLT4 LIGAND) (FLT4-			
DE	L)			
GN	VEGFC.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_Taxid:9606;			
RN	[1]			
RX	SEQUENCE FROM N.A., AND SEQUENCE OF 103-120.			
RX	MEDLINE-96178224; PubMed-8617204;			
RA	Joukov V., Pajusola K., Kaipainen A., Chillov D., Lahtinen I., Kuk E.,			
RA	Saksela O., Kalkkinen N., Alitalo K.;			
RT	"A novel vascular endothelial growth factor, VEGF-C, is a ligand for			
RT	the Flt4 (VEGFR-3) and KDR (VEGFR-2) receptor tyrosine kinases.";			
RL	EMBO J. 15:290-298(1996).			
RN	[2]			
RX	MEDLINE-96203094; PubMed-8612600;			
RA	Joukov V., Pajusola K., Kaipainen A., Chillov D., Lahtinen I., Kuk E.,			
RA	Saksela O., Kalkkinen N., Alitalo K.;			
RT	EMBO J. 15:1751-1751(1996).			
RN	[3]			
RX	SEQUENCE FROM N.A.			
RA	MEDLINE-96312526; PubMed-8700872;			
RA	Lee J., Gray A., Yuan J., Luo S.-M., Avraham H., Wood W.I.;			
RT	"Vascular endothelial growth factor-related protein: a ligand and			
RT	specific activator of the tyrosine kinase receptor Flt4.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 93:1988-1992(1996).			
RN	[4]			
RX	SEQUENCE FROM N.A.			
RA	Fitz L., Morris J.C., Towler P.S., Long A.J., Greco R.,			
RA	Burgess P., Giannotti J., Charette A., Hennessey D., Kovacic S.,			
RA	Fitzgerald M., Scalfire H., Weich N., Neden S., Finnelly H.,			
RA	Zollner R., Wang J., Nickbarg E., Gassaway R., Turner K.,			
RT	Wood C.R.;			
RT	Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.			
RT	- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL			
RT	CELL GROWTH.			
RT	- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.			
RT	- PTM: PROBABLY PROTEOLITICALLY PROCESSED IN THE C-TERMINUS.			
RT	- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.			
RT	-----			
RT	This SWISS-PROT entry is copyright. It is produced through a collaboration			
RT	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
RT	the European Bioinformatics Institute. There are no restrictions on its			
RT	use by non-profit institutions as long as its content is in no way			
RT	modified and this statement is not removed. Usage by and for commercial			
RT	entities requires a license agreement (See http://www.isb-sib.ch/announce/			

34	79	13.2	116	1	GLHA_ICTPU	O9ygp3
35	78.5	13.1	118	1	GLHA_CTEID	ictalurus p
36	78.5	13.1	118	1	GLHA_HYPMO	p3083 ctenopharynx
37	78.5	13.1	547	1	CO9_HORSE	p37037 hypopharynx
38	76	12.7	116	1	GLHA_CLAGA	p48770 equus caball
39	76	12.7	118	1	GLHA_CYPCA	p53542 claritas gar
40	74.5	12.5	118	1	GLHA_BOVIN	p18857 cyprinus ca
41	74.5	12.5	120	1	GLHA_BOVIN	p01217 b glycoprot
42	74.5	12.5	120	1	GLHA_SHEEP	p01218 o glycoprot
43	73.5	12.3	117	1	GLHA_ACALA	p30970 acanthopagr
44	73	12.2	117	1	GLHA_MORSA	091119 morone saxa
45	72	12.0	120	1	GLHA_HORSE	p01220 e glycoprot

```

CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: X94216: CAA63907.1: -
DR EMBL: U43142: AAB85214.1: -
DR EMBL: U58111: AAB02909.1: -
DR HSSP: P15692: 1VPF.
DR MIM: 601528: -
DR InterPro: IPR000072: -
DR InterPro: IPR002400: -
DR Pfam: PF00341: PDGF_1.
DR PRINTS: PR00438: GFCYSKNOT.
DR PROSITE: PS00249: PDGF_1; 1.
DR PROSITE: PS0278: PDGF_2; 1.
DR Mitogen: Growth factor; Glycoprotein; Signal; Repeat.
KW SIGNAL
FT PROPEP 1 102 POTENTIAL.
FT CHAIN 103 419 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
FT DOMAIN 275 365 4 X 24 AA TANDEM REPEATS.
FT REPEAT 275 298 1.
FT REPEAT 299 322 2.
FT REPEAT 323 346 3.
FT REPEAT 347 365 4 (PARTIAL).
FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 419 AA: 46883 MM: 9F598719D83E014F CRC64;

Query Match 63.0%; Score 377; DB 1; Length 419;
Best Local Similarity 63.6%; Pred. No. 8.7e-32;
Matches 63; Conservative 15; Mismatches 21; Indels 0; Gaps 0;

QY 2 YDIETLKVIDEWMORTQCSPRETCVEVASELGRKSTNFFKPCVNVFRCGCCNEESLIC 61
   1: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 114 YNFEILKSIDNEMRKTCQMPREVCIDVGKEFGVATNFFKPCVSVYRGGCCNSEGLQC 173
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 62 MNTSTYSISKOLFESVPLTSVPELVVYKVAHNTGCKL 100
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 174 MNTSTYLSKTLFEITVPLSGRKPVTISFANHSCRCM 212
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
VEGF_MOUSE STANDARD; PRT; 415 AA.
ID VEGF_MOUSE
AC P97953;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR (VEGF-C) (FLT4 LIGAND)
DE (FLT4-L).
GN VEGFC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=97388482; PubMed=9247316;
RA Filtz L.J., Morris J.C., Towler P., Long A., Burgess P., Greco R.,
RA Wang J., Gassaway R., Nickbarg E., Kovacic S., Charette A.,
RA Giamotti J., Finerty H., Zollner R., Beier D.R., Leak L.V.,
RA Turner K.J., Wood C.R.;
RL "Characterization of murine Flt4 ligand/VEGF-C.";


```

```

RL Oncogene 15:613-618(1997).
CC -1- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
CC CELL GROWTH.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: U73620: AAC52984.1: -
DR EMBL: U58112: AAB46707.1: -
DR HSSP: P15692: 1VPF.
DR MGI: 109124: Vegfc.
DR InterPro: IPR000072: -
DR InterPro: IPR002400: -
DR Pfam: PF00341: PDGF_1.
DR PRINTS: PR00438: GFCYSKNOT.
DR PROSITE: PS00249: PDGF_1; 1.
DR PROSITE: PS0278: PDGF_2; 1.
DR Mitogen: Growth factor; Glycoprotein; Signal; Repeat.
KW SIGNAL
FT PROPEP 1 98 POTENTIAL.
FT CHAIN 99 415 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
FT DOMAIN 271 361 4 X 24 AA TANDEM REPEATS.
FT REPEAT 271 294 1.
FT REPEAT 295 318 2.
FT REPEAT 319 342 3.
FT REPEAT 343 361 4 (PARTIAL).
FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 415 AA: 46471 MM: D9D3D53CECC659D6 CRC64;

Query Match 62.5%; Score 374; DB 1; Length 415;
Best Local Similarity 62.6%; Pred. No. 1.8e-31;
Matches 62; Conservative 15; Mismatches 22; Indels 0; Gaps 0;

QY 2 YDIETLKVIDEWMORTQCSPRETCVEVASELGRKSTNFFKPCVNVFRCGCCNEESLIC 61
   1: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 110 YNFEILKSIDNEMRKTCQMPREVCIDVGKEFGVATNFFKPCVSVYRGGCCNSEGLQC 169
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 62 MNTSTYSISKOLFESVPLTSVPELVVYKVAHNTGCKL 100
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 170 MNTSTYLSKTLFEITVPLSGRKPVTISFANHSCRCM 208
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
VEGF_MOUSE STANDARD; PRT; 214 AA.
ID VEGF_MOUSE
AC Q00731;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR PRECURSOR (VEGF) (VASCULAR
DE PERMEABILITY FACTOR) (VPF).
GN VEGF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92274860; PubMed=1592003;
RA Breiter G., Albrecht U., Sterner S., Rissu W.;
RT "Expression of vascular endothelial growth factor during embryonic
RT angiogenesis and endothelial cell differentiation.";
RL Development 114:521-532(1992).


```

```

RN [2]
RP SEQUENCE FROM N.A. (VEGF-1).
RA MEDLINE=9235593; PubMed=1644816;
RA Clafey K.P., Wikstrom W.O., Spiegelman B.M.;
RT "Vascular endothelial growth factor. Regulation by cell
RT differentiation and activated second messenger pathways.";
RL J. Biol. Chem. 267:16317-16322(1992).
RN [3]
RP SEQUENCE OF 1-3 FROM N.A.
RA MEDLINE=96216498; PubMed=8632007;
RA Shima D.T., Kuroki M., Deutsch U., Ng Y., Adams A.P., D'Amore P.A.;
RT "The mouse gene for vascular endothelial growth factor. Genomic
RT structure, definition of the transcriptional unit, and
RT characterization of transcriptional and post-transcriptional
RT regulatory sequences.";
RL J. Biol. Chem. 271:3877-3883(1996).
CC -1- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
CC CELL GROWTH. INDUCES ENDOTHELIAL PROLIFERATION AND VASCULAR
CC PERMEABILITY.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: VEGF-1 AND VEGF-2 ARE SECRETED WHILE
CC VEGF-3 REMAINS CELL-SURFACE ASSOCIATED UNLESS RELEASED BY
CC HEPARIN.
CC -1- ALTERNATIVE PRODUCTS: THREE FORMS (VEGF-1, VEGF-2 AND VEGF-3) ARE
CC PRODUCED AS A RESULT OF ALTERNATIVE SPLICING OF THE SAME GENE. THE
CC LONGEST FORM (VEGF-3, SHOWN HERE) CONTAINS A BASIC INSERT LINKED
CC TO CELL-ASSOCIATION/HEPARIN-BINDING.
CC -1- TISSUE SPECIFICITY: IN DEVELOPING EMBRYOS, EXPRESSED MAINLY IN
CC THE CHOROID PLEXUS, PARAVENTRICULAR NEUROEPITHELIUM, PLACENTA AND
CC KIDNEY GLOMERULI. ALSO FOUND IN BRONCHIAL EPITHELIUM, ADRENAL
CC GLAND AND IN SEMINIFEROUS TUBULES OF TESTIS. HIGH EXPRESSION OF
CC VEGF CONTINUES IN KIDNEY GLOMERULI AND CHOROID PLEXUS IN ADULTS.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S37052; AAB2252.1; -
DR EMBL: S38083; AAB2253.1; -
DR EMBL: S38100; AAB2254.1; -
DR EMBL: M95200; AAA0547.1; -
DR EMBL: U41383; -, NOT_ANNOTATED_CDS.
DR PIR: A43351; A43351.
DR HSSP: P15692; 2VGH.
DR MGD: MGI:103178; VEGF.
DR InterPro: IPR000072; -.
DR Pfam: PF00341; PDGF_1;
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS50278; PDGF_2; 1.
KW Mitogen; Growth factor; Glycoprotein; Alternative splicing; Signal.
FT SIGNAL 1 26
FT CHAIN 27 214 VASCULAR ENDOTHELIAL GROWTH FACTOR.
FT DISULFID 51 93 BY SIMILARITY.
FT DISULFID 82 127 BY SIMILARITY.
FT DISULFID 86 129 BY SIMILARITY.
FT DISULFID 76 76 INTERCHAIN (BY SIMILARITY).
FT DISULFID 85 85 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 100 100 N-LINKED (GLCNAc... ) (PROBABLE).
FT VARSPLIC 140 140 K -> N (IN ISOFORM VEGF-1).
FT VARSPLIC 141 164 MISSING (IN ISOFORM VEGF-1).
FT VARSPLIC 141 208 MISSING (IN ISOFORM VEGF-2).
FT CONFLICT 117 118 GE -> ER (IN REF. 2).
SQ SEQUENCE 214 AA; 25283 MW; B5540B51E4BBE17 CRC64;

```

Query Match 29.0%; Score 173.5; DB 1; Length 214;
 Best Local Similarity 37.6%; Pred. No. 4.8e-11;
 Matches 38; Conservative 16; Mismatches 44; Indels 3; Gaps 3;

```

OY 5 ETLKVIDEWMQRTQCSRETCVAVASLCKSTNTFFKPCVNFRCGGCCNESLICMT 64
DB 38 EVIKFMD-VYQSYCRIFELVDIFQEPDEIEYITKPSVPLMRAGCCNDALCEVPT 96
OY 65 STSYISKOLEFISVPLTSPVLPVKVANTGCKCLPTAR 105
DB 97 SESNITMOMIRIK-PHOS-OHIGEMFLQHSRCECRPKKR 135

RESULT 4
VEGF RAT STANDARD; PRT; 190 AA.
AC P16612;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR PRECURSOR (VEGF) (VASCULAR
DE PERMEABILITY FACTOR) (VPF).
GN VEGF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-190.
RA MEDLINE=90207249; PubMed=2320579;
RA Conn G., Bayne M.L., Soderman D.D., Kwok P.W., Sullivan K.A.,
RA Palisi T.M., Hope D.A., Thomas K.A.;
RT "Amino acid and cDNA sequences of a vascular endothelial cell mitogen
RT that is homologous to platelet-derived growth factor.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:2628-2633(1990).
CC -1- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
CC CELL GROWTH. INDUCES ENDOTHELIAL PROLIFERATION AND VASCULAR
CC PERMEABILITY.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR
CC TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN (BY
CC SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE PITUITARY, IN BRAIN, IN
CC PARTICULARLY IN SUPRAOPTIC AND PARAVENTRICULAR NUCLEI AND THE
CC CHOROID PLEXUS. ALSO FOUND ABUNDANTLY IN THE CORPUS LUTEUM OF
CC THE OVARY AND IN KIDNEY GLOMERULI.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M32167; AAA41211.1; -
DR PIR: A35987; A35987.
DR HSSP: P15692; 2VGH.
DR InterPro: IPR000072; -.
DR Pfam: PF00341; PDGF_1;
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS50278; PDGF_2; 1.
KW Mitogen; Growth factor; Glycoprotein; Signal.
FT SIGNAL 1 26
FT CHAIN 27 190 VASCULAR ENDOTHELIAL GROWTH FACTOR.
FT DISULFID 51 93 BY SIMILARITY.
FT DISULFID 82 127 BY SIMILARITY.
FT DISULFID 86 129 BY SIMILARITY.
FT DISULFID 76 76 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 100 100 N-LINKED (GLCNAc... ).
SQ SEQUENCE 190 AA; 22396 MW; 58374010441F377 CRC64;

```

Query Match 28.7%; Score 171.5; DB 1; Length 190;

```

Best Local Similarity 36.6%, Pred. No. 6,9e-11;
Matches 37, Conservative 17, Mismatches 44, Indels 3, Gaps 3;

OY      5 ETLKVIDEMORTQCSPRETCVEVASELGSKTNIFFKPCPVNFRGCGCNESTLCMNT 64
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      38 EVVKFMD-VYORSYCRPIETLVLDIQEYPDELTLEYIFKPSCVPLMRACGCCNDLEALCVPT 96

OY      65 STSYISKOLFESVPLTVSPELVPKVANHHQCCKLPAPR 105
       ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      97 SESNVWQIMRIK-PHQ-SQHIGMSFTLSHRCCECPKKDR 135

RESULT      5
VEGF_HUMAN  STANDARD:      PRT:      215 AA.
AC          PI5692;
DT          01-APR-1990 (Rel. 14, Created)
DT          01-APR-1990 (Rel. 14, Last sequence update)
DT          15-JUL-1999 (Rel. 38, Last annotation update)
DE          VASCULAR ENDOTHELIAL GROWTH FACTOR PRECURSOR (VEGF) (VASCULAR
DE          PERMEABILITY FACTOR) (VPF).
DE          VEGF OR VEGFA.
OS          Homo sapiens (Human).
OC          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX          NCBI_TaxID=9606;
[1]
RP        SEQUENCE FROM N.A.
RX        MEDLINE=90069608; PubMed=2479986;
RA        Leung D.W., Cachianes G., Kuang W.-J., Goeddel D.V., Ferrara N.;
RT        "Vascular endothelial growth factor is a secreted angiogenic
RT        mitogen."
RL        Science 246:1306-1309(1989).
[2]
RP        SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX        MEDLINE=90069609; PubMed=2479987;
RA        Keck P.J., Hauser S.D., Krivi G., Sanzo K., Warren T., Feder J.,
RT        Connolly D.T.;
RT        PDGF.";
RL        "Vascular permeability factor, an endothelial cell mitogen related to
Science 246:1309-1312(1989).
[3]
RP        SEQUENCE FROM N.A.
RX        MEDLINE=91268072; PubMed=1711045;
RA        Tischer E., Mitchell R., Hartman T., Silva M., Gospodarowicz D.,
RA        Fiddes J.C., Abraham R.A.;
RT        "The human gene for vascular endothelial growth factor. Multiple
RT        protein forms are encoded through alternative exon splicing.";
RL        J. Biol. Chem. 266:11947-11954(1991).
[4]
RP        SEQUENCE FROM N.A.
RX        MEDLINE=92231879; PubMed=1567395;
RA        Weindel K., Marne D., Welch H.A.;
RT        "AIDS-associated Kaposi's sarcoma cells in culture express vascular
RT        endothelial growth factor.";
RL        Biochem. Biophys. Res. Commun. 183:1167-1174(1992).
[5]
RP        PRELIMINARY SEQUENCE OF 27-36; 43-50 AND 59-81.
RX        MEDLINE=90062112; PubMed=2584205;
RA        Connolly D.T., Olander J.V., Heuvelman D., Nelson R., Monsell R.,
RA        Siegel N., Haymore B.L., Lelmgriuber R., Feder J.;
RT        "Human vascular permeability factor. Isolation from U937 cells.";
RL        J. Biol. Chem. 264:20017-20024(1989).
[6]
RP        SEQUENCE OF 27-41.
RX        MEDLINE=93145946; PubMed=7678805;
RA        Fleibich B.L., Jaeger B., Schoellmann C., Weindel K., Wilting J.,
RA        Kochs G., Marne D., Hug H., Welch H.A.;
RT        "Synthesis and assembly of functionally active human vascular
RT        endothelial growth factor homodimers in insect cells.";
RL        Eur. J. Biochem. 211:19-26(1993).
[7]
RP        X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 34-135.

```

RX MEDLINE=97352774; PubMed=9207067
 RA Muller V.A., Li B., Christinger H.W., Wells J.A., Cunningham B.C.,
 de Vos A.M.;
 RT "Vascular endothelial growth factor: crystal structure and functional
 mapping of the kinase domain receptor binding site."
 RT Proc. Natl. Acad. Sci. U.S.A. 94:7192-7197(1997).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (1.93 ANGSTROMS) OF 34-135.
 RA MEDLINE=98035455; PubMed=9351807;
 RX Muller V.A., Christinger H.W., Keyt B.A., de Vos A.M.;
 RA "The crystal structure of vascular endothelial growth factor (VEGF)
 RT refined to 1.93-A resolution: multiple copy flexibility and receptor
 RT binding".
 RL Structure 5:1325-1338(1997).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 39-134.
 RX MEDLINE=9119204; PubMed=9922142;
 RA Wiseman C., Christinger H.W., Cochran A.G., Cunningham B.C.,
 Fairbrother W.J., Keenan C.J., Meng G., de Vos A.M.;
 RA "Crystal structure of the complex between VEGF and a receptor-blocking
 RT peptide".
 RT Biochemistry 37:17765-17772(1998).
 RN [10]
 RP STRUCTURE BY NMR OF 34-135.
 RX MEDLINE=97477915; PubMed=9336848;
 RA Fairbrother W.J., Champe M.A., Christinger H.W., Keyt B.A.,
 Starovasnik M.A.;
 RA "Solution structure of the heparin-binding domain of vascular
 RT endothelial growth factor".
 RL Structure 6:637-648(1998).
 CC -I- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
 CC CELL GROWTH. INDUCES ENDOTHELIAL PROLIFERATION AND VASCULAR
 CC PERMEABILITY.
 CC -I- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -I- SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR
 CC TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN (BY
 CC SIMILARITY).
 CC -I- ALTERNATIVE PRODUCTS: FOUR FORMS OF VEGF ARE PRODUCED BY
 CC ALTERNATIVE SPLICING OF THE SAME GENE (VEGF-121, VEGF-165,
 CC VEGF-189 AND VEGF-215).
 CC -I- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC
 CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
 CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
 CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON
 CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
 CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
 CC ENTITIES REQUIRES A LICENSE AGREEMENT (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M32977; AAA35789.1; -;
 DR EMBL; M27281; AAA36807.1; -;
 DR EMBL; M63978; AAA36804.1; -;
 DR EMBL; M63971; AAA36804.1; JOINED.
 DR EMBL; M63972; AAA36804.1; JOINED.
 DR EMBL; M63973; AAA36804.1; JOINED.
 DR EMBL; M63974; AAA36804.1; JOINED.
 DR EMBL; M63975; AAA36804.1; JOINED.
 DR EMBL; M63976; AAA36804.1; JOINED.
 DR EMBL; M63977; AAA36804.1; JOINED.
 DR EMBL; X62568; CAA44447.1; -;
 DR PIR; A34492; A34492.
 DR PIR; A40079; A40079.
 DR PIR; A40080; A40080.
 DR PIR; A40454; A40454.

RESULT	6			
VEGF_SHEEP				
ID	VEGF_SHEEP	STANDARD;	PRT;	146 AA.
AC	p50412;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	VASCULAR ENOTHELIAL GROWTH FACTOR PRECURSOR (VEGF) (VASCULAR PERMEABILITY FACTOR) (VPF).			
DE	VEGF.			
GN				
OS	Ovis aries (Sheep).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Caprinae; Ovis.			
OX	NCBI_TaxID=9940;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Kidney;			
RX	MEDLINE=97117958; PubMed=8958842;			
RA	Reidner D.A., Dai Y., Li J., Charnock-Jones D.S., Smith S.K.,			
RA	Reynolds L.P., Moor R.M.;			
RT	"Characterization and expression of vascular endothelial growth			
RT	factor (VEGF) in the ovine corpus luteum."			
RL	J. Reprod. Fertl. 108:157-165(1996).			
CC	-1- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENOTHELIAL			
CC	CELL GROWTH. INDUCES ENOTHELIAL PROLIFERATION AND VASCULAR			
CC	PERMEABILITY.			
CC	-1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.			

RESULT	7			
ID	VEGF_BOVIN	STANDARD:	PRT:	190 AA.
AC	P15691;			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	VASCULAR ENDOTHelial GROWTH FACTOR PRECURSOR (VEGF) (VASCULAR PERMEABILITY FACTOR) (VPF).			
DE	VEGF.			
GN	Bos taurus (bovine).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 27-47.			
RX	MEDLINE=90069608; PubMed=2479986;			
RA	Leung D.W., Cachianes G., Kuang W.-J., Goeddel D.V., Ferrara N.;			
RT	"Vascular endothelial growth factor is a secreted angiogenic			
RL	mitogen.";			
RL	Science 246:1306-1309,(1989).			
RP	[2]			
RP	SEQUENCE OF 27-190 FROM N.A.			
RX	MEDLINE=90121225; PubMed=2610687;			
RA	Tischer E., Gospodarowicz D., Mitchell R., Silva M., Schilling J.,			
RA	Lau K., Crisp T., Fiddes J.C., Abraham J.A.;			
RT	"Vascular endothelial growth factor: a new member of the platelet-			
RT	derived growth factor gene family.";			

```

RL Biochem. Biophys. Res. Commun. 165:1198-1206(1989).
RN [3]
RP SEQUENCE OF 27-31.
RX MEDLINE=89286596; PubMed=2735925;
RA Perrera N., Henzel W.J.;
RT Pituitary follicular cells secrete a novel heparin-binding growth
RT factor specific for vascular endothelial cells.
RL Biochem. Biophys. Res. Commun. 161:851-858(1989).
CC -1- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
CC CELL GROWTH. INDUCES ENDOTHELIAL PROLIFERATION AND VASCULAR
CC PERMEABILITY.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR
CC TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M32976; AAA30502.1; -
DR EMBL: M31836; AAA30804.1; -
DR EMBL: M33750; AAA30805.1; -
DR PIR: A33255; A33255.
DR PIR: A33787; A33787.
DR PIR: B40080; B40080.
DR HSSP: P15692; 2VGH.
DR InterPro: IPR000072; -.
DR Pfam: PF00341; PDGF_1; 1.
DR PROSITE: PS00249; PDGF_2; 1.
DR PROSITE: PS0278; PDGF_2; 1.
KW Mitogen; Growth factor; Glycoprotein; Alternative splicing; Signal.
FT SIGNAL 1 26
FT CHAIN 1 27
FT DISULFID 51 93
FT DISULFID 82 127
FT DISULFID 86 129
FT DISULFID 76 76
FT DISULFID 85 85
FT CARBOHYD 100 100
FT VARSLIC 139 183
FT VARSLIC 184 184
SQ SEQUENCE 190 AA; 22310 MW; EDBF903E46E24789 CRC64;

Query Match 27.8%; Score 166.5; DB 1; Length 190;
Best Local Similarity 38.1%; Pred. No. 2.3e-10;
Matches 37; Conservative 14; Mismatches 43; Indels 3; Gaps 3;

```

```

OC Viruses; dsDNA-viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Parapoxvirus.
OX NCBI_TaxID=73495;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94076465; PubMed=8254780;
RA Lytle D.J., Fraser K.M., Fleming S.B., Mercer A.A., Robinson A.J.;
RT "Homologs of vascular endothelial growth factor are encoded by the
RT poxvirus of virus."
RL Virol. 68:84-92(1994).
CC -1- FUNCTION: INDUCES ENDOTHELIAL PROLIFERATION.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S67522; AAB29223.1; -
DR HSSP: P15692; 1VPE.
DR InterPro: IPR000072; -.
DR Pfam: PF00341; PDGF_1; FALSE_NEG.
DR PROSITE: PS00249; PDGF_1; FALSE_NEG.
DR PROSITE: PS0278; PDGF_2; 1.
KW Mitogen; Growth factor; Glycoprotein; Signal.
FT SIGNAL 1 148
FT CHAIN 1 148
FT DISULFID 46 88
FT DISULFID 77 130
FT DISULFID 81 132
FT DISULFID 71 71
FT DISULFID 80 80
FT CARBOHYD 95 95
SQ SEQUENCE 148 AA; 16078 MW; F0E13BA104CC73F8 CRC64;

Query Match 27.4%; Score 164; DB 1; Length 148;
Best Local Similarity 32.8%; Pred. No. 3.2e-10;
Matches 38; Conservative 12; Mismatches 40; Indels 26; Gaps 4;

```

```

OY 5 ETLKVIDEMORTQCSPRETCVEVASSELGKSTNFFKPPCVNFRGCGCNEESLICMNT 64
DB 38 EYVKFMD-VYORSFCRPIETLVDIFOEYDELEIFKPPSCVPLMRGCGCNDSELCVPT 96
OY 65 STYSIKOLFELISVPLTSPVPLPVKVNHTGCKLP 101
DB 97 EEPNITMOTMRK-PROS-OHIGEMSFLDHNNKCECRP 131

RESULT 8
VEGF_ORFN7 STANDARD; PRT; 148 AA.
AC P52585;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR HOMOLOG PRECURSOR.
GN A2R.
OS Off virus (strain NZ7) (OV NZ-7).

```

```

OY 69 ISKOLFELISVPLTSPVPLPVKVNHTGCKLP 107
DB 96 TV-----VAVSVTVGSSSGTNGSVSTNLRISVTEHTKDCDIGNTTTPTTRER 146

RESULT 9
VEGF_PIG STANDARD; PRT; 190 AA.
AC P49151;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR PRECURSOR (VEGF) (VASCULAR
DE PERMEABILITY FACTOR) (VPF).
GN VEGF.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=95143284; PubMed=7841203;
RX Sharma H.S., Tang Z.H., Gho B.C.H., Verdouw P.D.;

```

RT "Nucleotide sequence and expression of the porcine vascular
 RT endothelial growth factor."
 RL Biochim. Biophys. Acta 1260:235-238(1995).
 CC -1- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
 CC CELL GROWTH. INDUCES ENDOTHELIAL PROLIFERATION AND VASCULAR
 CC PERMEABILITY (BY SIMILARITY).
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR
 CC TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN (BY
 CC SIMILARITY). BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X81380; CAA57143.1; -
 DR HSSP: P15692; 2VGH.
 DR InterPro: IPR000072; -
 DR Pfam: PF00341; PDGF_1; 1.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS50278; PDGF_2; 1.
 KM Mitogen; Growth factor; Glycoprotein; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 190
 FT DISULFID 51 93
 FT DISULFID 82 127
 FT DISULFID 86 129
 FT DISULFID 76 76
 FT DISULFID 85 85
 FT CARBOHYD 100 100
 SQ SEQUENCE 190 AA; 22368 MW; 04D408BD7913047F CRC64;
 Query Match 27.3%; Score 163.5; DB 1; Length 190;
 Best Local Similarity 34.3%; Pred. No. 4.6e-10;
 Matches 36; Conservative 14; Mismatches 44; Indels 11; Gaps 3;
 QY 5 ETLKVIDEMQRTQSCPRETCVAVSELGKSTMTFFKPCVAVNFCGCCNEESLTCMNT 64
 DB 38 EYVKFMD-VYGRSYCRPIETLVDIPEYDEIYFRPSCVPLMRGCGCCNDDEGLECVPT 96
 QY 65 STYISKQLEFISVPLTSVP---ELVPKVVANHTGCKCLPTAPR 105
 DB 97 EEFNITWQIMRIK-----PHOGHIGEMSFLLQHNKCECRPKRDR 135
 RESULT 10
 PLGF_HUMAN STANDARD; PRT; 170 AA.
 AC P49763;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PLACENTA GROWTH FACTOR PRECURSOR (PLGF-1/PLGF-2).
 GN PGF OR PLGF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A. (PLGF-1).
 RC TISSUE-Placenta;
 RX MEDLINE=92021031; PubMed=1974389;
 RA Maglione D., Guerriero G., Vigiiletto G., Delli-Bovi P., Persico M.G.;
 RT "Isolation of a human placenta cDNA coding for a protein related to
 RT the vascular permeability factor."
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9267-9271(1991).
 [2]

RP SEQUENCE FROM N.A. (PLGF-2).
 RC TISSUE-Placenta;
 RX MEDLINE=94198032; PubMed=8148155;
 RA Hauser S., Welch H.A.;
 RT "A heparin-binding form of placenta growth factor (PLGF-2) is
 RT expressed in human umbilical vein endothelial cells and in
 RT placenta."
 RL Growth factors 9:259-268(1993).
 RN [3]
 RP SEQUENCE FROM N.A. (PLGF-2)
 RX MEDLINE=93205407; PubMed=7681160;
 RA Maglione D., Guerriero V., Vigiiletto G., Ferraro M.G., Aprelikova O.,
 RA Allitto K., del Vecchio S., Lei K.-J., Chou J.Y., Persico M.G.;
 RT "Two alternative mRNAs coding for the angiogenic factor, placenta
 RT growth factor (PLGF), are transcribed from a single gene of
 RT chromosome 14."
 RL Oncogene 8:925-931(1993).
 RN [4]
 RP CHARACTERIZATION, AND SEQUENCE OF 19-24.
 RX MEDLINE=95014370; PubMed=7929268;
 RA Park J.E., Chen H.H., Winer J., Houck K.A., Ferrara N.;
 RT "Placenta growth factor. Potentiation of vascular endothelial growth
 RT factor bioactivity, in vitro and in vivo, and high affinity binding
 RT to Flt-1 but not to Flk-1/KDR."
 RL J. Biol. Chem. 269:25646-25654(1994).
 CC -1- FUNCTION: GROWTH FACTOR OF UNKNOWN FUNCTION. BINDS TO RECEPTOR
 CC VEGFR-1 (FLT1). THE LONGER FORM (PLGF-2) CAN ALSO BIND HEPARIN. IT
 CC IS ABLE TO POTENTIATE THE ACTION OF LOW LEVELS OF VEGF.
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -1- SUBCELLULAR LOCATION: BOTH FORMS ARE SECRETED BUT THE LONGER FORM
 CC APPEARS TO REMAIN CELL ATTACHED UNLESS RELEASED BY HEPARIN.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: PLGF-1 AND PLGF-2 (SHOWN HERE);
 CC ARE PRODUCED BY ALTERNATIVE SPLICING. PLGF-1 DIFFERS FROM PLGF-2
 CC IN LACKING A 21 RESIDUES SEGMENT IN THE C-TERMINAL SECTION WHICH
 CC ACTS AS A CELL RETENTION SIGNAL.
 CC -1- TISSUE SPECIFICITY: WHILE BOTH FORMS ARE PRESENT IN MOST PLACENTAL
 CC TISSUES, THE LONGER FORM IS SPECIFIC TO EARLY (8 WEEK) PLACENTA
 CC AND ONLY THE SHORTER FORM IS FOUND IN THE COLON AND MAMMARY
 CC CARCINOMAS.
 CC -1- PTM: N-GLYCOSYLATED.
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X54936; CAA38698.1; -
 DR EMBL: A18411; CAA01393.1; -
 DR EMBL: S72960; AAB30462.2; -
 DR HSSP: P15692; 1VPF.
 DR MIM: 601121; -
 DR InterPro: IPR000072; -
 DR Pfam: PF00341; PDGF_1; 1.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS50278; PDGF_2; 1.
 KM Mitogen; Growth factor; Glycoprotein; Signal; Alternative splicing;
 KW Heparin-binding.
 FT SIGNAL 1 18
 FT CHAIN 19 170
 FT DISULFID 52 94
 FT DISULFID 83 128
 FT DISULFID 87 130
 FT DISULFID 77 77
 FT DISULFID 86 86
 FT CARBOHYD 33 33
 FT CARBOHYD 101 101
 FT VARSPLIC 142 162
 FT CONFLICT 91 91
 SQ SEQUENCE 170 AA; 19325 MW; E47639AC59C0963F CRC64;
 D -> N (IN REF. 1).
 MISSING (IN ISOFORM PLGF-1).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 INTERCHAIN (BY SIMILARITY).
 INTERCHAIN (BY SIMILARITY).
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 PLACENTA GROWTH FACTOR.
 BY SIMILARITY.
 BY SIMILARITY.

Query Match 26.6%; Score 159; DB 1; Length 170;
Best Local Similarity 34.4%; Pred. No. 1.2e-09;
Matches 31; Conservative 13; Mismatches 44; Indels 2; Gaps 1;

OY 12 EEMORTQCSPRETCVEVASLKGSTNTFFKPPCVNFRGCGCCNEESLJCMNTSTYSISK 71
DB 45 EVMGRSYCRALERTLVDSYSEVHEMFSPSCVSLRGTGCGGDDDLHCVPETANVM 104
OY 72 QFEISVPLTSPVLPVAVNHTGCKCLP 101
DB 105 QLKLR-SGDRPSVELFESQHVRCERCP 132

RESULT 11
VEGF_CAVPO STANDARD; PRT; 164 AA.

AC P26617;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR (VEGF) (VASCULAR PERMEABILITY FACTOR) (VEF).
GN VEGF.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Hystriocognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RA Beise B.;
RL Submitted (XXX-1992) to the EMBL/Genbank/DBJ databases.
CC -!- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL CELL GROWTH. INDUCES ENDOTHELIAL PROLIFERATION AND VASCULAR PERMEABILITY.
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -!- SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M84230; AAA37057.1; -.
DR HSSP; P15692; 2VGH.
DR InterPro; IPR000072; -.
DR Pfam; PF00341; PDGF_1; 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS50278; PDGF_2; 1.
KW Mitogen; Growth factor; Glycoprotein.
FT DISULFID 25 67 BY SIMILARITY.
FT DISULFID 56 101 BY SIMILARITY.
FT DISULFID 60 103 BY SIMILARITY.
FT DISULFID 50 50 INTERCHAIN (BY SIMILARITY).
FT DISULFID 59 59 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 74 74 N-LINKED (GLCNAc...) (POTENTIAL).
SQ SEQUENCE 164 AA; 19330 MW; 9EB86A81A9D5DCA4 CRC64;

Query Match 26.0%; Score 155.5; DB 1; Length 164;
Best Local Similarity 33.7%; Pred. No. 2.7e-09;
Matches 34; Conservative 16; Mismatches 40; Indels 11; Gaps 3;

OY 5 ETLKVIDEEMORTQCSPRETCVEVASLKGSTNTFFKPPCVNFRGCGCCNEESLJCMNT 64
DB 12 EEWKFMQ-VYKRSYCRPLEMLVDIFOEYPPDELEYLFKPSCVPLMRGCGCCNDESLCVPT 70

OY 65 STYSISKOLFESVPLTSPV-----ELVPVAVNHTGCKCLP 101
DB 71 EEFNTMQLMRK-----PHQGHIQEMSLFHSHKCECRP 105

RESULT 12
VEGF_ORFN2 STANDARD; PRT; 133 AA.

AC P52584;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR HOMOLOG PRECURSOR.
GN A2R.
OS Orf virus (strain NZ2) (OV NZ-2).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae; Parapoxvirus.
OX NCBI_TaxID=10259;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94076465; PubMed=8254780;
RA Lytle D.J., Fraser K.M., Fleming S.B., Mercer A.A., Robinson A.J.;
RT "Homologs of vascular endothelial growth factor are encoded by the Poxvirus orf virus".
RT J. Virol. 68:84-92(1994).
CC -!- FUNCTION: INDUCES ENDOTHELIAL PROLIFERATION.
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC -----
CC EMBL; S67520; AAB29220.2; -.
DR HSSP; P15692; 1VPE.
DR InterPro; IPR000072; -.
DR Pfam; PF00341; PDGF_1; 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS50278; PDGF_2; 1.
KW Mitogen; Growth factor; Glycoprotein; Signal.
FT SIGNAL 1 133 POTENTIAL.
FT CHAIN 1 133 VASCULAR ENDOTHELIAL GROWTH FACTOR
FT DISULFID 36 78 BY SIMILARITY.
FT DISULFID 67 112 BY SIMILARITY.
FT DISULFID 71 114 BY SIMILARITY.
FT DISULFID 61 61 INTERCHAIN (BY SIMILARITY).
FT DISULFID 70 70 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 85 85 N-LINKED (GLCNAc...) (POTENTIAL).
SQ SEQUENCE 133 AA; 14715 MW; 917C0F6883030C39 CRC64;

Query Match 24.7%; Score 147.5; DB 1; Length 133;
Best Local Similarity 35.2%; Pred. No. 1.5e-08;
Matches 32; Conservative 13; Mismatches 41; Indels 5; Gaps 2;

OY 17 TQCSPRETCVEVASLKGSTNTFFKPPCVNFRGCGCCNEESLJCMNTSTYSISKOLF 76
DB 34 SECKRPRIYVPVSETHPELTSGRFNPPCVTLMRGCGCCNDESLCVPEEVAVNSHLLGA 93
OY 77 SVPLTSPVLPVAVNHTGCKCLP---TAP 104
DB 94 SG--SGSNQMQLSFEVHKKCDRCRPFRTTP 122

RESULT 13
VEGF_MOUSE STANDARD; PRT; 188 AA.
ID VEGF_MOUSE

AC P49765;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR B PRECURSOR (VEGF-B) (VASCULAR
ENDOTHELIAL GROWTH FACTOR RELATED PROTEIN) (VRF).
GN VEGF OR VRF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=96197355; PubMed=8637916;
RA Olofsson B., Pajusola K., Kaipainen A., von Euler G., Joukov V.,
Saksela O., Orpana A., Pettersson R.F., Alltalo K., Eriksson U.,
RT "Vascular endothelial growth factor B, a novel growth factor for
endothelial cells.";
RT Proc. Natl. Acad. Sci. U.S.A. 93:2576-2581(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96183052; PubMed=8607868;
RA Tomson S., Lagercrantz J., Grimmond S., Silins G.,
RA Nordenskjöld M., Weber G., Hayward N.K.;
RT "Characterization of the murine VEGF-related factor gene.";
RT Biochem. Biophys. Res. Commun. 220:922-928(1996).
CC -1- FUNCTION: GROWTH FACTOR FOR ENDOTHELIAL CELLS. BINDS HEPARIN.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED. CAN ALSO FORM HETERODIMER
WITH VEGF.
CC -1- SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR
TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN.
CC -1- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN HEART, BRAIN, KIDNEY
AND SKELETAL MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U48800; AAB06273.1; -
CC EMBL: U43837; AAC52553.1; -
CC DR HSSP: P15692; 2VGH.
CC DR MGD: MGI:106199; Vegfb.
CC DR InterPro: IPR000072; -
CC DR Pfam: PF00341; PDGF_1;
CC DR PROSITE: PS00249; PDGF_1; 1.
CC DR PROSITE: PS50278; PDGF_2; 1.
CC KW Mitogen; Growth factor; Signal; Heparin-binding.
CC FT SIGNAL 1 21
CC FT CHAIN 1 21
CC FT CHAIN 22 188
CC FT CHAIN 188 AA; 21442 MW; D52A055FB955E9CA CRC64;
SQ SEQUENCE 188 AA; 21442 MW; D52A055FB955E9CA CRC64;
Query Match 24.7%; Score 147.5; DB 1; Length 188;
Best Local Similarity 33.0%; Pred. No. 2, 1e-08;
Matches 29; Conservative 15; Mismatches 41; Indels 3; Gaps 1;
OY 14 MORTQSPRETCVEVASLSEKSNFTFFKPPCVNVFRCGCGCNEESLIQNTSISYISKOL 73
DB 42 YAAATQCPREVVVPLSMELMGVNVKQLVPSCVTVGRGCGCPDDGLECVPTGHOVROMOI 101
OY 74 FEISVPLTSVPELVKPVKVNHTGCKCLP 101
DB 102 LMTQYPSQOLGEM---SLEHSGCECRP 126
RESULT 14

VEGF_HUMAN
ID VEGF_HUMAN STANDARD; PRT; 188 AA.
AC P49765;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR B PRECURSOR (VEGF-B) (VEGF RELATED
FACTOR).
GN VEGF OR VRF.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=96197355; PubMed=8637916;
RA Olofsson B., Pajusola K., Kaipainen A., von Euler G., Joukov V.,
Saksela O., Orpana A., Pettersson R.F., Alltalo K., Eriksson U.,
RT "Vascular endothelial growth factor B, a novel growth factor for
endothelial cells.";
RT Proc. Natl. Acad. Sci. U.S.A. 93:2576-2581(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97077124; PubMed=8919691;
RA Grimmond S., Lagercrantz J., Drinkwater C., Silins G., Tomson S.,
RA Pollock P., Gotley D., Carson E., Rakar S., Nordenskjöld M., Ward L.,
RA Hayward N., Weber G.;
RT "Cloning and characterization of a novel human gene related to
vascular endothelial growth factor.";
RT Genome Res. 6:124-131(1996).
CC -1- FUNCTION: GROWTH FACTOR FOR ENDOTHELIAL CELLS. BINDS HEPARIN.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED. CAN ALSO FORM HETERODIMER
WITH VEGF.
CC -1- SUBCELLULAR LOCATION: SECRETED BUT REMAINS ASSOCIATED TO CELLS OR
TO THE EXTRACELLULAR MATRIX UNLESS RELEASED BY HEPARIN.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXCEPT LIVER.
CC HIGHEST LEVELS FOUND IN HEART, SKELETAL MUSCLE AND PANCREAS.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U48801; AAB06274.1; -
CC EMBL: U43369; AAA91463.1; -
CC DR HSSP: P15692; 1VPE.
CC DR MIM: 601398; -
CC DR InterPro: IPR000072; -
CC DR Pfam: PF00341; PDGF_1;
CC DR PROSITE: PS00249; PDGF_1; 1.
CC DR PROSITE: PS50278; PDGF_2; 1.
CC KW Mitogen; Growth factor; Signal; Heparin-binding.
CC FT SIGNAL 1 21
CC FT CHAIN 1 21
CC FT CHAIN 22 188
CC FT CHAIN 188 AA; 21261 MW; F04654D5A3727194 CRC64;
SQ SEQUENCE 188 AA; 21261 MW; F04654D5A3727194 CRC64;
Query Match 24.5%; Score 146.5; DB 1; Length 188;
Best Local Similarity 32.3%; Pred. No. 2, 6e-08;
Matches 31; Conservative 16; Mismatches 44; Indels 5; Gaps 2;
OY 8 KVID--EEMORTQSPRETCVEVASLSEKSNFTFFKPPCVNVFRCGCGCNEESLIQNTS 65
DB 34 KVVSWIDVYTRATQCPREVVVPLVELMGVAVAKQLVPSCVTVGRGCGCPDDGLECVPTG 93
OY 66 TSYISKOLEFISVPLTSVPELVKPVKVNHTGCKCLP 101
DB 94 OHOVROMOILMIRYPSQOLGEM---SLEHSGCECRP 126

RESULT 15
PLGF_MOUSE
ID PLGF_MOUSE STANDARD; PRT; 158 AA.
AC P49764;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PLACENTA GROWTH FACTOR PRECURSOR (PLGF).
GN PGF OR PLGF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart; N.A.
RX MEDLINE=97059399; PubMed=8903720;
RA Dipalma T., Tucci M., Russo G., Maglione D., Lago C.T., Romano A.,
RA Saccone S., della Valle G., de Gregorio L., Dragani T.A.,
RA Vigiiletto G., Persico M.G.;
RT "The placenta growth factor gene of the mouse."
RL Mamm. Genome 7:6-12(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NIH SWISS;
RX MEDLINE=98065381; PubMed=9401819;
RA Achen M.G., Gad J.M., Stacker S.A., Wilks A.F.;
RT "Placenta growth factor and vascular endothelial growth factor are
co-expressed during early embryonic development."
RL Growth Factors 15:69-80(1997).
CC -!- FUNCTION: GROWTH FACTOR ACTIVE IN ANGIOGENESIS, AND ENDOTHELIAL
CELL GROWTH (BY SIMILARITY).
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
CC or send an email to license@isb-sdb.ch).

CC
DR EMBL: X80171; CAA56453.1; -
DR EMBL: X96793; CAA65587.1; -
DR HSSP: P15692; 1VPR.
DR MGD: MGI:105095; Pgf.
DR InterPro: IPR000072; -.
DR Pfam: PF00341; PDGF_1; 1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS0278; PDGF_2; 1.
KW Mitogen; Growth factor; Glycoprotein; Signal.
FT SIGNAL 1 18
FT CHAIN 19 158
FT DISULFID 48 90
FT DISULFID 79 125
FT DISULFID 83 127
FT DISULFID 73 73
FT DISULFID 82 82
FT CARBOHYD 29 29
FT CARBOHYD 30 30
FT CARBOHYD 97 97
SQ SEQUENCE 158 AA; 17876 MW; F16128BEA0790438 CRC64;

Query Match 23.5%; Score 140.5; DB 1; Length 158;
Best Local Similarity 31.5%; Pred. No. 9, 1e-08;
Matches 29; Conservative 15; Mismatches 45; Indels 3; Gaps 2;

QY 11 DEMQRTGCSFRRTCYVASSELGKSTNTPFKPCVAVNFRGGCCGNEESLICANTSTSYIS 70
DB 40 NEWMGRSYCRPMKELVYIIIDEPYDEVSHIFSPCVLLSRCSCGCCGDBGHCVPIKNTANIT 99

QY 71 KOLFETSVPLTSVPEL-VPVKVANHTGCKCLP 101
DB 100 MQI--LKIPNRPDHPHYVENTFSQDVLCGCRP 129

Search completed: October 17, 2001, 14:51:47
Job time: 473 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2001, 14:50:56 ; Search time 128.06 Seconds
(without alignments)
112.613 Million cell updates/sec

Title: US-09-427-657-4_COPY_93_201
Perfect score: 598
Sequence: 1 FYDIETLKVIDEWQRTQCS.....KVANHGGCKLPAPRHPYS 109

Scoring table:
BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_16:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organelle:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-unclassified:*
13: sp-vertebrate:*
14: sp-virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	598	100.0	354	4	043915
2	579	96.8	326	11	035251
3	575	96.2	358	11	P97946
4	370	61.9	418	13	057352
5	369	61.7	420	6	09XS50
6	207	34.6	126	11	035757
7	178	29.8	122	6	09GLX1
8	172.5	28.8	190	11	09OX39
9	171.5	28.7	170	11	09JXK7
10	171.5	28.7	214	11	09OXG7
11	169.5	28.3	147	4	09UH58
12	169.5	28.3	171	4	09H1W8
13	169.5	28.3	174	4	09UL23
14	169.5	28.3	209	4	060720
15	169.5	28.3	232	4	09H1W9
16	169.5	28.3	254	4	016889
17	167.5	28.0	118	6	09MZB1
18	167.5	28.0	146	11	09OXG6
19	167.5	28.0	190	6	077643

20	167	27.9	142	11	09ERL6	09erl6 mesocricetu
21	166.5	27.8	190	6	09GKR0	09gkr0 equus caball
22	164.5	27.5	190	6	09XSF3	09xsf3 canis faml
23	164.5	27.5	208	6	09XSF4	09xsf4 canis faml
24	164.5	27.5	214	6	09XSF5	09xsf5 canis faml
25	164.5	27.5	214	6	09MYV3	09myv3 canis faml
26	163.5	27.3	190	6	09GLS2	09gl52 sus scrofa
27	163	27.3	148	13	042571	042571 xenopus lae
28	163	27.3	194	13	042572	042572 xenopus lae
29	162	27.1	132	14	09YMF3	09ymf3 orf virus.
30	161	26.9	149	6	09XS47	09xs47 bos taurus
31	160	26.8	149	4	09Y6S8	09y6s8 homo sapien
32	159.5	26.7	191	4	075875	075875 homo sapien
33	159	26.6	124	6	09GK00	09gk00 callithrix
34	155.5	26.0	110	11	088911	088911 rattus norv
35	154.5	25.8	141	11	070123	070123 mus musculu
36	147.5	24.7	207	11	064290	064290 mus musculu
37	146.5	24.5	188	6	09XS48	09xs48 bos taurus
38	146.5	24.5	193	6	09XS49	09xs49 bos taurus
39	146.5	24.5	207	4	016528	016528 homo sapien
40	145.5	24.3	158	11	063434	063434 rattus norv
41	144.5	24.2	150	11	054881	054881 rattus norv
42	139	23.2	68	6	097500	097500 oryctolagus
43	139	23.2	144	13	073822	073822 brachydanio
44	139	23.2	188	13	073682	073682 brachydanio
45	126	21.1	75	6	018843	018843 oryctolagus

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	354 AA.
043915				
ID	043915			
AC	043915			
DT	01-JUN-1998 (TREMBLrel. 06, Created)			
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)			
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)			
DE	GROWTH FACTOR FIGF.			
GN	FIGF OR VEGF-D.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_Taxid=9606;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98140120; PubMed=9479493;			
RA	Rocchigliani M., Lestingi M., Luddi A., Orlandini M., Franco B.,			
RA	Rossi A., Ballabio A., Zuffardi O., Oliviero S.;			
RT	"Human FIGF: cloning, gene structure, and mapping to chromosome Xp22.1			
RT	between the FIGA and the GPR genes.";			
RL	Genomics 47:207-216(1998).			
RP	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=LUNG;			
RX	MEDLINE=97349118; PubMed=9205122;			
RA	Yamada Y., Nezu J., Shimane M., Hirata Y.;			
RT	"Molecular cloning of a novel vascular endothelial growth factor,			
RT	VEGF-D.";			
RL	Genomics 42:483-488(1997).			
RP	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98118549; PubMed=9435229;			
RA	Achen M.G., Jeltsch M., Kukr E., Maekinen T., Vitale A., Wilks A.F.,			
RA	Allitalo K., Stacker S.A.;			
RT	"Vascular endothelial growth factor D (VEGF-D) is a ligand for the			
RT	tyrosine kinases VEGF receptor 2 (Flk1) and VEGF receptor 3 (Flt4).";			
RL	Proc. Natl. Acad. Sci. U.S.A. 95:548-553(1998).			
DR	EMBL: Y12864; CAA73371.1; JOINED.			
DR	EMBL: Y12865; CAA73371.1; JOINED.			
DR	EMBL: Y12866; CAA73371.1; JOINED.			
DR	EMBL: Y12867; CAA73371.1; JOINED.			
DR	EMBL: Y12868; CAA73371.1; JOINED.			

DR EMBL: Y12869; CAA73371.1; JOINED.
 DR EMBL: Y12870; CAA73371.1; JOINED.
 DR EMBL: D89630; BAA24264.1; -.
 DR EMBL: AJ000185; CAA03942.1; -.
 DR EMBL: Y12863; CAA73370.1; -.
 DR HSSP: P15692; 1VP.
 DR InterPro: IPR000072; -.
 DR Pfam: PF00341; PDGF; 1.
 DR ProDom: PD001629; -; 1.
 DR ProSITE: PS00249; PDGF_1; 1.
 DR ProSITE: PS50278; PDGF_2; 1.
 DR SMART: SM00141; PDGF; 1.
 SQ SEQUENCE 354 AA; 40444 MW; 2048D769D735173E CRC64;

Query Match 100.0%; Score 598; DB 4; Length 354;
 Best Local Similarity 100.0%; Pred. No. 1.7e-59;
 Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYDIETLKVIDEEMORTQCSPRETCVEVASELGSNTFFKPCPVNVFRGCGCNEESLI 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 93 FYDIETLKVIDEEMORTQCSPRETCVEVASELGSNTFFKPCPVNVFRGCGCNEESLI 152
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 61 CMMSTSYISKOLFELISVPLTSPELVPKVNHTGCKCLPTAPRHPYS 109
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 153 CMMSTSYISKOLFELISVPLTSPELVPKVNHTGCKCLPTAPRHPYS 201
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 2
 035251 PRELIMINARY; PRT; 326 AA.
 AC 035251;
 DT 01-JUN-1998 (TREMBLrel. 05, Created)
 DT 01-JUN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR D.
 GN VEGF-D.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE DAWLEY;
 RA MEDLINE=97349118; PubMed=9205122;
 RA Yamada Y., Nezu J., Shimane M., Hirata Y.;
 RT "Molecular cloning of a novel vascular endothelial growth factor,
 VEGF-D";
 RT Genomics 42:483-488(1997).
 RL EMBL: AF014827; AAB6557.1; -.
 DR HSSP: P15692; 1VP.
 DR InterPro: IPR000072; -.
 DR Pfam: PF00341; PDGF; 1.
 DR ProDom: PD001629; -; 1.
 DR ProSITE: PS00249; PDGF_1; 1.
 DR ProSITE: PS50278; PDGF_2; 1.
 DR SMART: SM00141; PDGF; 1.
 SQ SEQUENCE 326 AA; 37112 MW; 1261AFA373596C00 CRC64;

Query Match 96.8%; Score 579; DB 11; Length 326;
 Best Local Similarity 94.5%; Pred. No. 2.2e-57;
 Matches 103; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 FYDIETLKVIDEEMORTQCSPRETCVEVASELGSNTFFKPCPVNVFRGCGCNEESLI 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 98 FYDIETLKVIDEEMORTQCSPRETCVEVASELGSNTFFKPCPVNVFRGCGCNEESLY 157
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 61 CMMSTSYISKOLFELISVPLTSPELVPKVNHTGCKCLPTAPRHPYS 109
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 158 CMMSTSYISKOLFELISVPLTSPELVPKVNHTGCKCLPTAPRHPYS 206
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 3
 P97946 PRELIMINARY; PRT; 358 AA.
 AC P97946;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR D (C-FOS INDUCED GROWTH FACTOR).
 GN VEGF-D OR FIGF.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=97030254; PubMed=8876195;
 RA Ordandini M., Marconcini L., Ferruzzi R., Oliviero S.;
 RT "Identification of a c-fos-induced gene that is related to the
 platelet-derived growth factor/vascular endothelial growth factor
 family";
 RT Proc. Natl. Acad. Sci. U.S.A. 93:11675-11675(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG;
 RX MEDLINE=97349118; PubMed=9205122;
 RA Yamada Y., Nezu J., Shimane M., Hirata Y.;
 RT "Molecular cloning of a novel vascular endothelial growth factor,
 VEGF-D";
 RT Genomics 42:483-488(1997).
 RL EMBL: X99572; CAA67892.1; -.
 DR EMBL: D89628; BAA14002.1; -.
 DR HSSP: P15692; 1VP.
 DR MGI: 108037; Figf.
 DR InterPro: IPR000072; -.
 DR Pfam: PF00341; PDGF; 1.
 DR ProDom: PD001629; -; 1.
 DR ProSITE: PS00249; PDGF_1; 1.
 DR ProSITE: PS50278; PDGF_2; 1.
 DR SMART: SM00141; PDGF; 1.
 SQ SEQUENCE 358 AA; 40908 MW; 6636B17FBF07037C CRC64;

Query Match 96.2%; Score 575; DB 11; Length 358;
 Best Local Similarity 93.6%; Pred. No. 6.9e-57;
 Matches 102; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 FYDIETLKVIDEEMORTQCSPRETCVEVASELGSNTFFKPCPVNVFRGCGCNEESLI 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 98 FYDIETLKVIDEEMORTQCSPRETCVEVASELGSNTFFKPCPVNVFRGCGCNEGYM 157
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 61 CMMSTSYISKOLFELISVPLTSPELVPKVNHTGCKCLPTAPRHPYS 109
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 158 CMMSTSYISKOLFELISVPLTSPELVPKVNHTGCKCLPTAPRHPYS 206
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 4
 057352 PRELIMINARY; PRT; 418 AA.
 ID 057352;
 AC 057352;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR.
 GN VEGF-C.
 OS Coturnix coturnix japonica (Japanese quail).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Coturnix.
 RX NCBI_TaxID=93934;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98167900; PubMed=9435294;

RA Eichmann A., Corbel C., Jaffredo T., Breant V., Joukov V., Kumar V.,
RA Allitalo K., Le Douarin N.M.;
RT "Avian VEGF-C: cloning, embryonic expression pattern and stimulation
of the differentiation of VEGFR-expressing endothelial cell
precursors";
RL Development 125:743-752(1998).
DR EMBL: Y15837; CAA75799.1; -.
DR HSSP: P15692; IVP.
DR InterPro: IPR000072; -.
DR InterPro: IPR002400; -.
DR Pfam: PF00341; PDGF; 1.
DR PRINTS: PR00438; GFCYSKNOT.
DR PRODOM: PD001629; -; 1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS02278; PDGF_2; 1.
DR SMART: SM00141; PDGF; 1.
KM Signal.
FT SIGNAL.
FT CHAIN 111 418 POTENTIAL.
SQ SEQUENCE 418 AA; 46839 MW; 099BFCG79151BF2B CRC64;

Query Match 61.9%; Score 370; DB 13; Length 418;
Best Local Similarity 63.6%; Pred. No. 1e-33;
Matches 63; Conservative 14; Mismatches 22; Indels 0; Gaps 0;

OY 2 YDIETLKVIDEEMORQCSPRETCVEVASLKGSTNFFKPCVNVPRCGCCNEESLIC 61
DB 113 YNEILKSIDTEMRKTCQMPREVCVDGKEFGATNFFKPCVSTYRCGGCCNSEGLQC 172
OY 62 MNTSTYISKOLFELISVPLTSVPELVPKVANHGGKCL 100
DB 173 MNSTNISTKLTLEITVPLSGPKPVTYVSFANHSCRCM 211

RESULT 5
O9XS50 PRELIMINARY; PRT; 420 AA.
AC O9XS50;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEART;
RA Liu X., Yonekura H., Yamagishi S., Yamamoto Y., Yamamoto H.;
RT "Structure and expression of bovine VEGF family.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB004275; BAA77687.1; -.
DR HSSP: P15692; IVP.
DR InterPro: IPR000072; -.
DR Pfam: PF00341; PDGF; 1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS02278; PDGF_2; 1.
DR SMART: SM00141; PDGF; 1.
KM Signal.
FT SIGNAL.
FT CHAIN 21 20 POTENTIAL.
SQ SEQUENCE 420 AA; 46681 MW; 58BA84317A3CB2D CRC64;

Query Match 61.7%; Score 369; DB 6; Length 420;
Best Local Similarity 61.6%; Pred. No. 1.4e-33;
Matches 61; Conservative 16; Mismatches 22; Indels 0; Gaps 0;
OY 2 YDIETLKVIDEEMORQCSPRETCVEVASLKGSTNFFKPCVNVPRCGCCNEESLIC 61
DB 113 YNEILKSIDTEMRKTCQMPREVCVDGKEFGATNFFKPCVSTYRCGGCCNSEGLQC 172

DB 115 YNEILKSIDNEMRKTCQMPREVCIDGKEFGATNFFKPCVSVYRCGGCCNSEGQC 174
OY 62 MNTSTYISKOLFELISVPLTSVPELVPKVANHGGKCL 100
DB 175 MNSTYISKLTLEITVPLSGPKPVTYVSFANHSCRCM 213

RESULT 6
O35757 PRELIMINARY; PRT; 126 AA.
AC O35757;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR-C (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LUNG;
RA Mandriota S.J., Pepper M.S.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF010302; AAB63248.1; -.
DR HSSP: P15692; 2VPF.
DR InterPro: IPR000072; -.
DR Pfam: PF00341; PDGF; 1.
DR PROSITE: PS02278; PDGF_2; 1.
DR SMART: SM00141; PDGF; 1.
FT NON_TER 1 1
FT NON_TER 126 126
SQ SEQUENCE 126 AA; 13977 MW; 8F365AFBCA4E037B0 CRC64;

Query Match 34.6%; Score 207; DB 11; Length 126;
Best Local Similarity 62.5%; Pred. No. 7.6e-16;
Matches 35; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

OY 45 VNVPRCGCCNEESLICMNTSTYISKOLFELISVPLTSVPELVPKVANHGGKCL 100
DB 1 VSVYRCGGCCNSEGLQCMNTSTYISKLTLEITVPLSGPKPVTYVSFANHSCRCM 56
RESULT 7
O9GLX1 PRELIMINARY; PRT; 122 AA.
AC O9GLX1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR-D (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Mandriota S.J., Pepper M.S.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF099135; AAG29747.1; -.
FT NON_TER 1 1
FT NON_TER 122 122
SQ SEQUENCE 122 AA; 13820 MW; CC504B00E29D54EB CRC64;

Query Match 29.8%; Score 178; DB 6; Length 122;
Best Local Similarity 94.1%; Pred. No. 1.4e-12;
Matches 32; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 76 ISVPLTSVPELVKPVKVNHTGCKCLPTAPRHPYS 109
 Db 1 ISVPLTSVPELVKPVKVNHTGCKCLPTAPRHPYS 34

RESULT 8

Q90X39 PRELIMINARY; PRT; 190 AA.
 AC 090X39; 13, Created
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR.
 GN VEGF.
 OS Spalax leucodon ehrenbergi (Ehrenberg's mole rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Spalacinae;
 OC Spalax.
 NCBI_TaxID=30637;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99313148; PubMed=10386577;
 RA Avivi A., Resnick M.B., Nevo E., Joel A., Levy A.P.;
 RT "Adaptive hypoxic tolerance in the subterranean mole rat Spalax
 ehrenbergi: the role of vascular endothelial growth factor.";
 RL EMBL: 452133-140(1999).
 DR EMBL; AF186236; AAD56245.1; -
 DR HSSP; P15692; 2VPF.
 DR InterPro: IPR000072; -
 DR Pfam: PF00341; PDGF_1; 1.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS50278; PDGF_2; 1.
 DR SMART; SM00141; PDGF; 1.
 SQ SEQUENCE 190 AA; 2248 MW; 2228383BC65F0BFE CRC64;

Query Match 28.8%; Score 172.5; DB 11; Length 190;
 Best Local Similarity 37.6%; Pred. No. 8.8e-12;

Matches 38; Conservative 16; Mismatches 44; Indels 3; Gaps 3;

QY 5 ETLKVIDEEMORTQCSPRETCVEVASSELGKSTNTFFKPCVNVFRCGCCNEESLTCMNT 64
 Db 38 EYVKFMD-VYRSTCRPIETLVDFQYDELEYIFKPCSVPLMRACGCCNDALCEVPT 96
 QY 65 STSYISKQLEISVPLTSVPELVKPVKVNHTGCKCLPTAPR 105
 Db 97 SESNVTMQIMRIK-PHOS-OHIGEMSFLOHNRCECRPKRDR 135

RESULT 9

Q90XK7 PRELIMINARY; PRT; 170 AA.
 AC 090XK7;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR-A144.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ishii H., Arakawa T., Okayama M., Oota I., Takuma T., Inomata K.;
 RT "Developmental expression of vascular endothelial growth factor-A
 (VEGF-A) splicing variants, VEGF-A188, VEGF-A164, VEGF-A144, and VEGF-
 A120 in rat mesenter muscle.";
 RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF222779; AAF25958.1; -
 DR InterPro: IPR000072; -
 DR Pfam: PF00341; PDGF_1; 1.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS50278; PDGF_2; 1.
 DR PROSITE: PS50278; PDGF_2; 1.

DR SMART; SM00141; PDGF; 1.
 SQ SEQUENCE 170 AA; 20018 MW; B371BF4A67E3231B CRC64;

Query Match 28.7%; Score 171.5; DB 11; Length 170;
 Best Local Similarity 36.6%; Pred. No. 1e-11;
 Matches 37; Conservative 17; Mismatches 44; Indels 3; Gaps 3;

QY 5 ETLKVIDEEMORTQCSPRETCVEVASSELGKSTNTFFKPCVNVFRCGCCNEESLTCMNT 64
 Db 38 EYVKFMD-VYRSTCRPIETLVDFQYDELEYIFKPCSVPLMRACGCCNDALCEVPT 96
 QY 65 STSYISKQLEISVPLTSVPELVKPVKVNHTGCKCLPTAPR 105
 Db 97 SESNVTMQIMRIK-PHOS-OHIGEMSFLOHNRCECRPKRDR 135

RESULT 10

Q90XG7 PRELIMINARY; PRT; 214 AA.
 AC 090XG7;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR-A188.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ishii H., Arakawa T., Okayama M., Oota I., Takuma T., Inomata K.;
 RT "Developmental expression of vascular endothelial growth factor-A
 (VEGF-A) splicing variants, VEGF-A188, VEGF-A164, and VEGF-A120 in rat
 mesenter muscle.";
 RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF215725; AAF19211.1; -
 DR HSSP; P15692; 1VPP.
 DR InterPro: IPR000072; -
 DR Pfam: PF00341; PDGF; 1.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS50278; PDGF_2; 1.
 DR SMART; SM00141; PDGF; 1.
 SQ SEQUENCE 214 AA; 25239 MW; 60FBB876F5304946 CRC64;

Query Match 28.7%; Score 171.5; DB 11; Length 214;
 Best Local Similarity 36.6%; Pred. No. 1.3e-11;
 Matches 37; Conservative 17; Mismatches 44; Indels 3; Gaps 3;

QY 5 ETLKVIDEEMORTQCSPRETCVEVASSELGKSTNTFFKPCVNVFRCGCCNEESLTCMNT 64
 Db 38 EYVKFMD-VYRSTCRPIETLVDFQYDELEYIFKPCSVPLMRACGCCNDALCEVPT 96
 QY 65 STSYISKQLEISVPLTSVPELVKPVKVNHTGCKCLPTAPR 105
 Db 97 SESNVTMQIMRIK-PHOS-OHIGEMSFLOHNRCECRPKRDR 135

RESULT 11

Q90H58 PRELIMINARY; PRT; 147 AA.
 AC 090H58;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR ISOFORM 121 PRECURSOR
 (DJ261G23.6.2).
 GN VEGF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;

